Beyond Instrumental Value:
How Complexity Shapes Information Demand

Menglong Guan	Ryan Oprea	Sevgi Yuksel
Penn State	UC Berkeley	NYU

November 3, 2026

Abstract

We use novel methods to experimentally investigate how characteristics of information structures beyond instrumental value influence information demand. While subjects strongly favor structures with higher instrumental value, they systematically deviate from optimal choices in predictable ways. Participants display aversion to larger structures, such as those that generate more signals with residual uncertainty, appear to simplify their assessment of value by neglecting signal likelihoods, and respond less to value differences when comparisons are difficult. The structure characteristics behind these patterns are as predictive of behavior as individual characteristics such as cognitive ability, and are more predictive of errors among subjects who are otherwise more likely to act optimally. Our findings suggest that these distortions stem from the cognitive difficulty of evaluating information structures, with implications for models of information demand and the design of information provision.

We would like to thank numerous seminar participants for helpful comments. Guan: mkg6150@psu.edu; Oprea: roprea@gmail.com; Yuksel: sevgiy@gmail.com. This research was supported by the National Science Foundation under Grant SES-1949366 and was approved by the Institutional Review Boards of UC Santa Barbara and UC Berkeley. This paper replaces and subsumes our earlier working paper titled "Too Much Information".

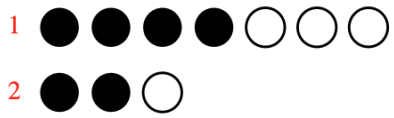
1 Introduction

Few products are chosen more frequently by modern consumers than sources of information. From news sources to social media influencers, from product reviewers to scientific experts, decision makers are continuously tasked with deciding what competing sources of information to rely on. How do people make these decisions? Economic theory provides a sharp answer: demand for sources of information should be shaped by instrumental value—the expected utility improvement the information can be expected to produce in relevant decisions. However, as a growing literature in economics suggests, assessing the value of complicated objects like information structures can be cognitively demanding, making instrumental value difficult to discern. As a result, characteristics of information structures other than instrumental value may significantly influence the way people make these choices.

In this paper, we report a novel experiment designed to better understand the role instrumental value plays in people's demand for information and to identify what other formal characteristics of information structures also shape this demand. To do this, we ask subjects to *rank* information structures drawn from a set of 237 structures that vary widely in a number of characteristics, including instrumental value. Higher-ranked structures are more likely to be later given to subjects to help them make an incentivized decision (a guess about the realization of a random variable), making this elicitation incentive compatible. We also collect information on subjects' ability to *use* signals from information structures exogenously assigned to them, allowing us to study the relationship between subjects' *demand* for information and their ability to optimally use it.

An important innovation in our design is a novel method for presenting information structures to subjects that we believe (and empirically verify) is particularly intuitive. States of the world are presented as a collection of events (black and white balls, shown on their screen), and information structures are presented as partitions of these events. Subjects receiving a signal from an information structure are simply told which cell of the partition the true event lies in (see Figure 1(a) for an example). This allows us to present information structures graphically to subjects, facilitating relatively fast assessment and enabling us to ask individual subjects to rank a number of information structures during the experiment. As we verify in the data, subjects generally make optimal use of signals under this representation, allowing us to study difficulties in evaluating information structures in a setting where errors in using their signals are rare.

As expected, we find that subjects' demand for (ranking of) information structures is strongly influenced by instrumental value. However, we also find that subjects frequently (in 27 percent of pairwise comparisons) rank less valuable structures above more valuable structures—an error



(a) Partition representation

Clue 1	Likelihood of Clue: 70%	Likelihood of Black: 57%; Likelihood of White: 43%
Clue 2	Likelihood of Clue: 30%	Likelihood of Black: 67%; Likelihood of White: 33%

(b) Tabular representation

Figure 1: Two Representations of the Same Information Structure *Notes: The subjects' task is to guess the color of a randomly selected ball. Panel (a) presents the information structure using the partition representation employed in the main study: the clue (signal) reveals whether the randomly selected ball is in group 1 or group 2. Panel (b) presents the same information structure using the tabular representation employed in the robustness experiment, showing the probability of each clue and the probability of each color conditional on the clue.*

that causes significant loss of earnings in the experiment. By contrast, conditional on receiving a signal from an information structure, subjects make near-perfect use of this information, making optimal guesses 94.6 percent of the time. This suggests that the demand errors we observe arise primarily from difficulties in ex-ante evaluation—assessing information structures and weighing tradeoffs between them—rather than from difficulties in optimally using their signals ex post.

Importantly, characteristics of information structures beyond instrumental value systematically predict which structures subjects prefer and when they make ranking errors. We examine a number of formal metrics (26 in total) of information structures and their comparisons and find that these characteristics together capture 82 percent of the explainable variation in preference rankings. Beyond their predictive power, these characteristics reveal three distinct patterns in how subjects evaluate information structures.

First, subjects are averse to information structures with a larger *size*. Size can reflect whether they induce more distinct posteriors, generate more signals with residual uncertainty (i.e., signals that are not fully revealing of the state), or simply have more signals overall. Because the size of an information structure affects how difficult it is to assess its instrumental value, we interpret this pattern as evidence of *complexity aversion* in information demand.

Second, subjects appear to simplify the task of assessing instrumental value. They prefer information structures that would have a higher instrumental value if (counterfactually) all signals were equally likely to occur. This unweighted measure of value predicts rankings over and above instrumental value itself. We interpret this as evidence that subjects use a natural simplifying heuristic to approximate the value of structures: they ignore, or insufficiently account for, signal probabilities, reducing the computational demands of assessing instrumental value.

Third, characteristics of the choice set as a whole, such as the total instrumental value of the two structures being compared and tradeoff complexity measures inspired by recent literature (Enke

& Shubatt 2026, Shubatt & Yang 2024), predict attenuation effects. Preference rankings are less responsive to the difference in instrumental value between two structures when their total value is low or when the tradeoffs between them are harder to assess, making them difficult to compare.

These patterns persist in an additional experiment that presents information structures in a more standard tabular format (Figure 1(b)) and elicits information demand through binary choices rather than preference rankings.[1] They also replicate in a laboratory experiment that uses the partition representation and includes treatment variations that explore mechanisms, as well as in an online replication of that experiment. Section 4.5 presents these additional studies and robustness analyses.[2]

To benchmark these results on information structure characteristics, we compare them to parallel findings on how characteristics of *decision makers* influence information demand. We find that structure characteristics influence information demand to a similar degree as, for instance, decision makers' cognitive test scores. Moreover, we find a surprising relationship between subject sophistication (e.g., cognitive scores) and responses to structure characteristics beyond instrumental value: these characteristics are *more* predictive of ranking preference among more sophisticated subjects, even though these subjects make fewer errors overall. This suggests that the errors of less sophisticated subjects tend to be more random (and less predictable), whereas those of more sophisticated subjects are more systematic and closely related to characteristics of the information structures. Additionally, we find that the structure characteristics are more predictive of behavior among ambiguity-averse subjects, suggesting a connection between responses to complexity and attitudes towards ambiguity (see de Clippel, Moscariello, Ortoleva & Rozen (2024) for related evidence).

We interpret these results as evidence that evaluating the relative value of information sources is a cognitively demanding task, even when the information they subsequently provide is easy to use. Comparing two information structures requires considering every signal that each of them can generate, identifying the best action after each signal, and aggregating the resulting payoffs by the likelihoods of the signals. Our findings show that people respond to these cognitive demands in systematic and multifaceted ways. As a result, characteristics of information structures beyond instrumental value influence how people choose between information sources, and the characteristics

[1]The partition design used in the main study indeed appears to facilitate optimal behavior: optimality is higher than in the robustness experiment, both for guesses (94.6 versus 90.1 percent) and for information demand (73.2 versus 67.7 percent), despite the simpler binary-choice elicitation used in the latter experiment.

[2]Furthermore, within the main study, the patterns hold among subjects who made no mistakes on comprehension questions and among those who never guessed suboptimally.

that predict choices are informative about the specific ways in which people cope with complexity. Importantly, these responses follow predictable patterns, which persist across subject pools, representations of information structures, and elicitation formats. The cognitive demands of evaluating and comparing information are therefore amenable to modeling. Our results characterize the regularities that such a model should account for to better explain and predict how people seek out and consume information.

Related Literature

Our paper contributes to several literatures.

Information Demand. First, we make a methodological contribution to the growing experimental literature studying information demand. Our design allows us to study demand for information in a setting where making optimal use of information is very easy and does not require difficult Bayesian reasoning. Thus, our visual representation of information structures successfully excludes non-Bayesian reasoning or misinterpretation of information structures as confounds for studying the demand for information. Similar experimental designs can be used to study a wide range of questions about people's taste for information.

Second, our paper adds to a growing literature studying how factors other than instrumental value influence information demand. Prior studies have experimentally or theoretically examined the role of confidence or confirmation seeking (Jones & Sugden 2001, Eliaz & Schotter 2007, 2010, Charness, Oprea & Yuksel 2021, Montanari & Nunnari 2022), preference for certainty (Ambuehl & Li 2018, Novák, Matveenko & Ravaioli 2024), timing of resolution of uncertainty (Grant, Kajii & Polak 1998, Nielsen 2020, Zimmermann 2015), skewness of information sources (Masatlioglu, Orhun & Raymond 2023), symmetry of information sources (Llorente-Saguer, Oliveros & Zultan 2025), anticipatory feelings (Caplin & Leahy 2001), motivated attention (Falk & Zimmermann 2024, Golman & Loewenstein 2018, Golman, Loewenstein, Molnar & Saccardo 2022), and behavioral motivations stimulated by changes in beliefs like disappointment aversion (Palacios-Huerta 1999, Dillenberger 2010, Andries & Haddad 2020), loss aversion (Koszegi & Rabin 2009), dissonance avoidance (Festinger 1957), suspense and surprise (Ely, Frankel & Kamenica 2015), etc., play in information demand.

Most closely related to our study in this literature is Liang (2023), a concurrent paper that studies suboptimal valuation of information structures and provides evidence that is broadly consistent with the mechanism underlying our main results. In particular, his results suggest that subjects have difficulty foreseeing and integrating payoffs from multiple information-contingent choices, but it is mostly difficulties with integration that get in the way of optimal valuation.

We contribute to this literature in several ways: (i) our experimental paradigm generates rich variation that enables us to examine how information demand responds to different structure characteristics; (ii) we compare the predictive power of these characteristics to individual-level characteristics such as cognitive ability; and (iii) we analyze interaction effects between structure characteristics and individual-level characteristics.

Complexity. Our study relates to a growing literature showing that cognitive limitations and complexity (i.e., costly information processing) play an important role in a wide range of economically important settings. For instance, recent work has provided evidence suggesting that complexity has an important influence on behavior in canonical settings like lottery choice (Enke & Graeber 2023, Oprea 2024), intertemporal decision making (Enke, Graeber & Oprea 2025) and Bayesian updating (Ba, Bohren & Imas 2023). See Oprea (2026) and Enke (2024) for recent reviews of this literature.

Perhaps most closely related to our paper in this literature is Enke & Shubatt (2026), which develops interpretable indices of complexity in lottery choice problems that are predictive of error rates in identifying the lottery with the highest expected value. We identify characteristics of information structures that are predictive of information demand and document systematic patterns of their impacts.[3,4]

Overall, we contribute to the literature on complexity by showing that information processing constraints extend to the no less canonical task of evaluating information. In this context, we document three patterns consistent with responses to complexity emphasized in prior work. First, we find evidence that decision makers are systematically *averse* to information structures that are complex in a specific, formal sense. Complexity aversion has been particularly emphasized in the context of lottery choice (Puri 2025), and recent work suggests it may be linked to ambiguity aversion arising from the complexity of the environment (de Clippel et al. 2024). Our findings are consistent with such a link. Second, we find that decision makers tend to *simplify* the evaluation of information structures by, for instance, neglecting the likelihoods of their signals, in line with the view that complexity leads decision makers to replace a demanding computation with a simpler

[3] An information structure can be mapped into a compound lottery. However, its value comes from how it helps improve decision making, so this mapping requires specifying a signal-contingent action rule and identifying its payoff consequences. A lottery, by contrast, directly specifies payoffs and their probabilities. Which characteristics predict choices between information structures therefore remains a question of its own.

[4] Beyond the difference in research questions, our approach also differs in several methodological respects. While Enke & Shubatt (2026) focus primarily on characteristics defined at the level of the choice set, we also examine differences between the two information structures, benchmark the predictive power of structure characteristics against individual-level characteristics, and study how responses to structure characteristics vary across individuals.

procedure (Enke 2024, Oprea 2026, Arrieta & Nielsen 2025). Third, we find that certain types of comparisons between information structures are complex for subjects in the sense that they produce *attenuation*—an inelastic response to fundamentals that has been associated with cognitive limitations in recent work (Enke, Graeber, Oprea & Yang 2026). A distinctive feature of our study is that we document all three effects—*aversion*, *heuristic simplification*, and *attenuation*—within a unified experimental framework and show how they coexist and shape behavior in complementary ways.

The remainder of the paper is organized as follows. Section 2 presents our theoretical framework. Section 3 describes the experimental design. Section 4 presents results. Section 5 concludes by discussing the implications of our results.

2 Theoretical Framework

2.1 Instrumental Value of Information

Consider a finite state space Ω, with a typical state denoted by ω. The prior distribution on Ω is denoted by p. An information structure, indexed by i, is a stochastic mapping from the state space Ω to a finite set of signals S. It is useful to think of i as inducing a distribution over posteriors.[5] That is, given p, an information structure i induces (i) a distribution q^i over S and (ii) conditional on each signal s, a posterior distribution p^i_s over the state space.

In standard economic theory, the value of an information structure to a decision-maker (DM) depends on the decision problem the information structure is meant to inform. Suppose the DM faces a decision problem in which she observes signal s from information structure i and takes action $a \in A$ to maximize $\mathbb{E}[u(a,\omega)|s] := \sum p^i_s(\omega)u(a,\omega)$, where $u(a,\omega)$ describes the decision-maker's state-dependent utility function. The *instrumental value* (or simply *value*) of i, given the set of actions A available and utility function u, is the expected increase in utility made possible by the DM being able to condition her action on the realized signals:

$$V^i = \underbrace{\sum_{s\in S} q^i(s) \underbrace{\max_{a\in A} \mathbb{E}[u(a,\omega)|s]}_{\substack{\text{Expected utility}\\ \text{conditional on } s}}}_{\text{Expectation over } s} - \underbrace{\max_{a\in A} \mathbb{E}[u(a,\omega)]}_{\substack{\text{Expected utility}\\ \text{without } s}}.$$

Instrumental value is the key characteristic shaping information demand (i.e., preferences for in-

[5] For each ω, let $\sigma^i_\omega(s) \in \Delta(S)$ denote the probability that signal s is produced by structure i. The probability of observing signal s is $q^i(s) := \sum_\omega p(\omega)\sigma^i_\omega(s)$. For each signal, the posterior distribution on Ω can be computed using Bayes' rule. Conditional on each signal s, $p^i_s(\omega) = \frac{p(\omega)\sigma^i_\omega(s)}{q^i(s)}$.

formation) in standard information economics.

2.2 A Guessing Task

We focus on a simple guessing task in our experiment to study the demand for information. The state of the world is binary, i.e. $\Omega = \{b, w\}$, where b (as will be clear in the next section) can be interpreted as referring to the color *black* and w referring to the color *white*. The decision-maker's prior is fixed at $p := p(b) = 0.6$. The decision-maker takes a binary action, $A = \{b, w\}$, with the goal of matching the state such that $u(a, \omega) = B$ (where B is a bonus) if $a = \omega$ and zero otherwise.

With this simple specification, the instrumental value of an information structure reduces to the following:

$$V^i = \Big(\underbrace{\sum_{s \in S} q^i(s) \underbrace{\max\{p^i_s, 1 - p^i_s\}}_{\substack{\text{Guessing accuracy} \\ \text{conditional on } s}}}_{\text{Expectation over } s} - \underbrace{p}_{\substack{\text{Guessing accuracy} \\ \text{without information}}} \Big) B, \tag{1}$$

where $q^i(s)$ is the probability that signal s is generated by structure i and p^i_s is the probability that $\omega = b$ conditional on signal s. Note that the expected utility of the decision-maker in this problem is equal to their guessing accuracy times the bonus associated with guessing correctly. Thus, the value of an information structure for a decision-maker, facing such a guessing task, is directly linked to the expected improvement in their guessing accuracy. In the following, we will simply refer to the expected improvement in guessing accuracy that a certain information structure induces as its instrumental value.

Two features of this task are worth noting. First, the optimal use of an information structure is intuitive: given a signal, the DM should guess the state that is more likely under the posterior, that is, b if $p^i_s > \frac{1}{2}$ and w if $p^i_s < \frac{1}{2}$. Second, the ranking of information structures by instrumental value does not depend on the DM's risk attitude. Since the DM receives either B or zero, a higher instrumental value simply corresponds to a higher probability of receiving B under optimal signal-contingent actions. Thus, any DM whose preferences are monotone in the probability of winning the bonus, including an expected-utility maximizer regardless of her risk attitude, would prefer the higher-value structure.[6]

Although this setting is simple, as we will show in the next section, it allows us to construct a rich set of information structures that differ not only in their instrumental value but also in a wide variety of characteristics that plausibly influence the information processing required to evaluate

[6]The argument, of course, does not cover preferences that violate the reduction of compound lotteries, such as preferences over the timing or the manner in which uncertainty is resolved (Kreps & Porteus 1978, Grant, Kajii & Polak 1998). We return to such preferences in Section 4.5.

information structures.

3 Experimental Design

This section describes the design of our main study. We discuss additional experiments that assess the robustness of our findings and explore their underlying mechanisms in Section 4.5.

3.1 Guessing Task and Representation of Information

Our experimental design is built around the simple guessing task introduced in the last section. A ball is randomly drawn from a set of 10 balls, six of which are always black and four of which are always white. The subjects' task is to correctly guess the color of that randomly selected ball. They earn a bonus of \$10 if they guess correctly and zero otherwise.

Before making their guesses, subjects receive information about the randomly selected ball from one information structure. Each information structure is represented as a partition of the 10 balls, i.e., a grouping of the 10 balls into non-empty cells. An information structure provides information about the randomly selected ball by revealing to the subject *which cell of the partition* the randomly selected ball belongs to. Thus, each cell of the partition can be interpreted as a distinct signal that the structure might generate. The size of a cell (the number of balls in it) visually represents the probability with which that signal will be realized, and the composition of the balls within the cell (relative share of black to white balls) visually represents the posterior conditional on the signal being realized. Presenting information in this way therefore makes the characteristics of an information structure particularly transparent to subjects.

Figure 1(a) provides an example of an information structure that partitions the 10 balls into two cells (two possible signals) displayed in two numbered rows. The first cell (labeled "1" and referred to as "group 1" in the experiment) consists of four black balls and three white balls, and the second cell ("group 2") consists of the remaining two black balls and one white ball. This information structure provides information about the randomly selected ball by revealing which cell (group 1 vs. group 2) the randomly selected ball belongs to. Formally, the presented information structure generates a binary signal. The first (second) signal, corresponding to group 1 (group 2), is realized with a 70 (30) percent chance. Conditional on the first (second) signal, the posterior probability that the color of the randomly selected ball is black is $\frac{4}{7}$ ($\frac{2}{3}$). The optimal guess is simply the majority color of the cell, which is black in both cases. In addition, as expressed in equation 1, these properties (the likelihood of each signal and the posterior conditional on this signal) are sufficient to determine the instrumental value of the information structure.

Using the partition representation, we construct *all* information structures with up to five

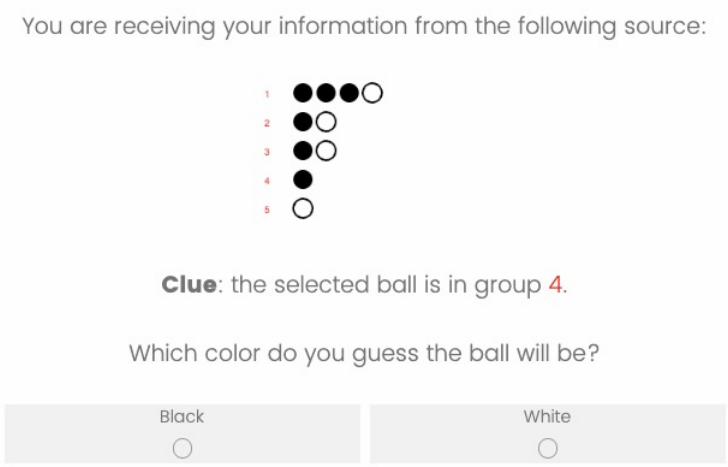


Figure 2: Guessing Task

signals. These correspond to all possible partitions of the 10 balls into one, two, three, four, and five cells, yielding 237 distinct structures in total. Our analysis focuses on pairwise comparisons between these structures. Theoretically, there are 27,966 possible pairwise comparisons. In our data, 23,859 of these are observed at least once, with some comparisons appearing up to 14 times. This induces substantial variation in instrumental value and also variation in a number of other structure characteristics, allowing us to explore how subjects respond to changes in instrumental value and other factors that characterize information structures.[7]

3.2 Eliciting Guesses

Part 1 of the experiment consists of fifteen *guessing* tasks. In each task, subjects guess the color of a randomly selected ball based on a signal generated from a given information structure. The 15 distinct information structures used across tasks are randomly selected at the subject level from the set of 237 structures, without replacement. Figure 2 shows a screenshot of a guessing task. Subjects observe a partition of the black and white balls, with the cells (groups) of the partition displayed in numbered rows. Subjects learn which group the selected ball is from and use this information to guess its color.

The purpose of the guessing tasks in the experiment is to study whether subjects can make optimal use of signals when they are presented using the partition design. This is important for interpreting their demand for information structures in Part 2.

[7]The complete list of structure characteristics considered is presented in Table 1 of Section 4.3.

3.3 Eliciting Preferences for Information Structures

Part 2 of the experiment elicits subjects' preferences for information structures using *ranking* tasks. In each task, subjects are asked to rank seven distinct information structures from most preferred to least preferred. In determining their rankings, subjects were incentivized to consider all possible pairwise comparisons between the seven structures presented to them. Specifically, subjects were told that two of the seven structures would be randomly selected, and they would receive information from the structure they had ranked higher. Our analysis accordingly focuses mainly on these pairwise comparisons.

Each subject is presented with eight ranking tasks. For each subject, the eight sets of seven structures (56 in total) used across these tasks are randomly selected without replacement from the 237 possible structures, subject to the constraint that each set includes structures spanning the full range of instrumental values.[8] This allows us to evaluate, in each ranking task, how demand for information responds to instrumental value and how this responsiveness is influenced by other characteristics of information structures that are not directly relevant to instrumental value (our primary question).

Figure 3 provides a screenshot of a ranking task. In the task, the seven selected information structures are presented in a randomized order and subjects are asked to reorder them using a drag-and-drop interface.[9]

3.4 Additional Tasks

The experiment also includes tasks designed to provide deeper insight into how individuals value information structures, as well as additional tasks that measure individual characteristics which we included to help us to interpret subjects' preferences for and use of information structures, allowing us to better assess heterogeneity across subjects.

At the end of Part 2, we ask subjects to indicate how confident they are that they ranked

[8]Given the prior $p = 0.6$, the baseline guessing accuracy without any additional information is 0.6. Thus, the instrumental value of an information structure, defined as the expected increase in accuracy it induces, can take on values of 0, 0.1, 0.2, 0.3, or 0.4 in our experiment. Each of the eight ranking tasks includes at least one information structure corresponding to each of these distinct values.

[9]We include seven information structures in each ranking task to elicit preferences over a substantial share of the 237 possible structures without overtaxing subjects. Without assuming transitivity, it would have required eliciting 21 pairwise choices to yield equivalent data to a ranking task. As discussed in Sections 4.2 and 4.5, the high optimality rate of rankings and the robustness of our findings across different elicitation formats, including binary choice, indicate that subjects found this elicitation manageable.

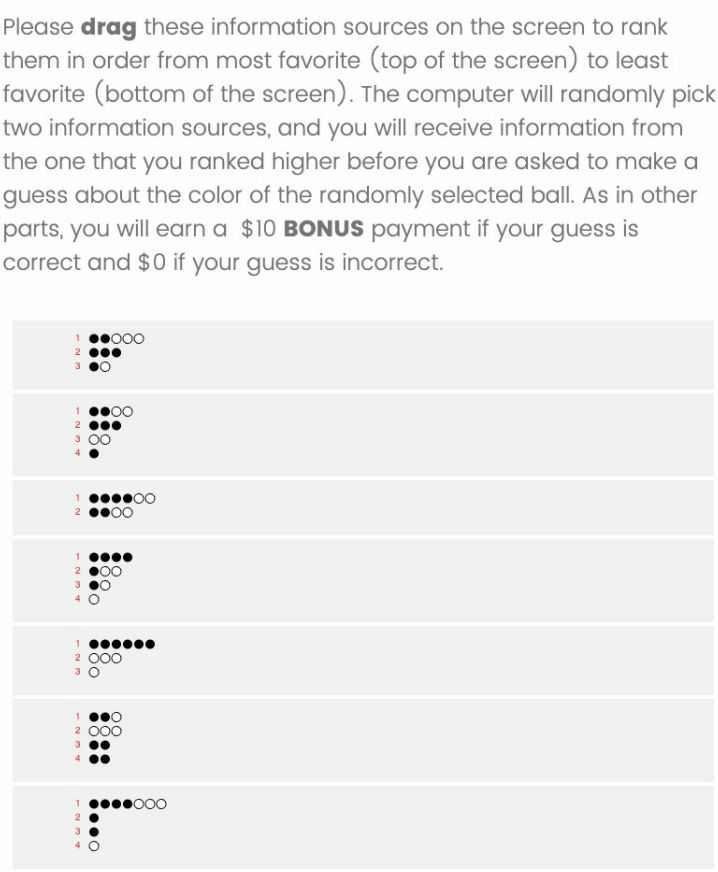


Figure 3: Ranking Task *Notes: As subjects drag and reorder structures, an automatically updating order number appears to the left of each structure. These numbers reflect the current ranking after each move. Check the experiment interface screenshots in Appendix F for an example.*

information structures in a way that maximizes their chance of winning the $10 bonus.[10] Part 3 of the experiment consists of the following tasks: (i) *Information structure computation*: subjects are asked to assess the chance that a computer would guess the color of a randomly selected ball correctly when it makes the optimal guess based on information from a given information structure, and report how confident they are that their assessments are correct;[11] (ii) *Risk and ambiguity attitude elicitation* (3 tasks): we use multiple price lists to elicit subjects' certainty equivalents for three lotteries with equal expected value, including a simple 50/50 lottery, a compound lottery,

[10]Specifically, we ask subjects: for any two information structures A and B where they ranked A higher than B, how confident they are that their chance of winning the $10 bonus (i.e., their guesses being correct) is as high or higher when receiving information from A versus B.

[11]This task provides a measure of subjects' ability to identify the instrumental value of an information structure. The information structure used in this task is one of two pre-selected information structures (randomized at the subject level) that have the same instrumental value but differ in other characteristics. The two structures can be found in the experiment interface screenshots in Appendix F.

and an Ellsberg lottery;[12] (iii) *Compound lottery computation*: subjects are asked to compute the exact chance that a certain outcome in the compound lottery is realized and indicate how confident they are about their computations.[13] Part 4 contains ten cognitive questions: one Wason selection task, three cognitive reflection test questions, three Raven's test questions, and three belief bias syllogism questions.[14] Details of all these additional tasks can be found in Appendix F.

3.5 Incentives and Implementation Details

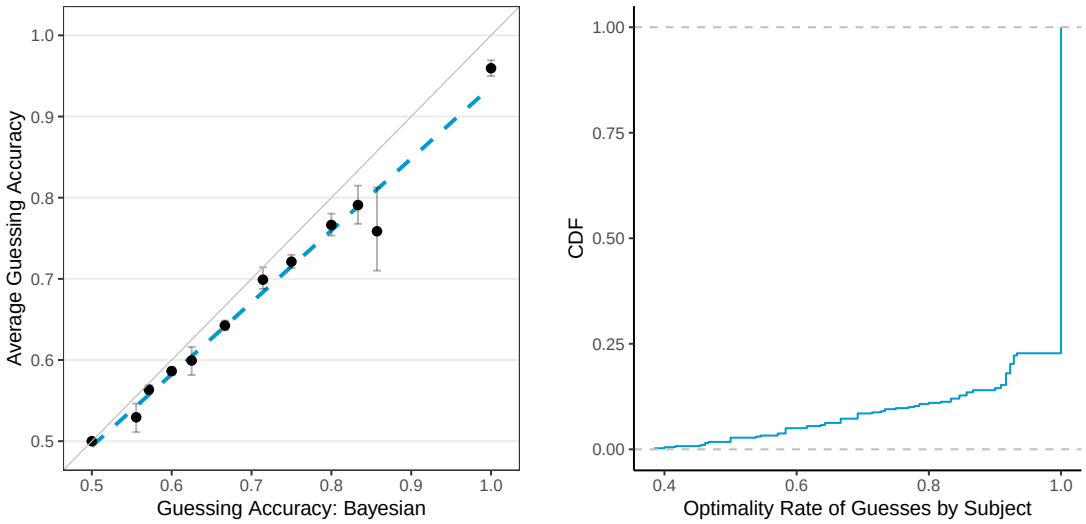
The experiment was conducted on Prolific in November 2024 with 400 participants, using software programmed by the authors in Qualtrics.

All participants received a base payment of \$6 and an extra \$0.5 if they answered all seven comprehension questions correctly on their first attempt. Comprehension questions are extensive and directly test participants' understanding of the partition design. In particular, participants are tested on the prior, the likelihood of different signal realizations from a structure and the posterior conditional on these signals (see Appendix F). We keep track of the number of mistakes each participant makes in the comprehension questions. We include all participants in the main analysis and discuss the robustness of our results to different subsamples in Section 4.5. A quarter of the participants were randomly selected to receive an additional bonus payment based on their decisions in one randomly selected part of the experiment. If part 1 was selected for payment, one of the 15 guessing tasks was randomly chosen and the participant received \$10 if they guessed the color of the selected ball correctly in that task. If part 2 was selected, the software randomly chose one ranking task and drew two information structures from that task. The participant was assigned to receive information from the higher-ranked structure in a final guessing task and received \$10 if guessing correctly in that task. If part 3 was selected, one randomly chosen task in the part determined the participant's bonus payment. If part 4 was selected, the participant received \$0.5 for each correctly answered cognitive question. The median time participants spent on the experiment was about 30 minutes and the average (median) bonus payment is about \$7.1 (\$5.5).

[12] The order of the three certainty equivalent elicitation tasks is randomized at the subject level. The tasks are taken from de Clippel et al. (2024).

[13] Both the information computation and compound lottery computation tasks are incentivized by the Binarized Scoring Rule (Hossain & Okui 2013). Inspired by Danz, Vesterlund & Wilson (2022), we do not explicitly present the detailed payoff calculation rules to subjects but allow them to access this information if they want.

[14] These questions are the same as those implemented in Oprea & Yuksel (2022).



(a) Guessing Accuracy by Value

(b) Distribution of Optimality of Guesses

Figure 4: Optimality of the Use of Signals *Notes: In panel (a), the best achievable (Bayesian) accuracy level depicts* $\max\{p_b^s, 1-p_b^s\}$ *where* p_b^s *is the posterior on the state being black given the signal* s*. Gray lines denote 95 percent confidence intervals by bootstrapping. The light blue dashed line is the best linear fit, and the gray solid line is the 45 degree line. In panel (b), the optimality of guesses is computed on the individual level and denotes the share of signals for which guess was optimal.*

4 Results

4.1 Optimality of Guesses

We begin by confirming that our experimental design is successful in mostly removing inferential errors like failures of Bayesian reasoning when subjects use signals from our information structures to inform their guesses. To do this, we study how subjects make use of the information provided to them. We focus on two measures. *Accuracy* is defined as the likelihood of a guess matching the state (the true color of the randomly selected ball). We contrast accuracy of a guess with the best achievable accuracy level given the signal realization. The *optimality* of a guess denotes whether the guess is in the direction (black or white) that maximizes accuracy given the signal realization.[15] To allow for clean interpretation of this measure, we restrict attention to signals with Bayesian posteriors different from 0.5 for which there is a unique optimal guess.

Figure 4a presents data on guessing accuracy, and shows how this varies with the best achievable (Bayesian) accuracy level given the signal. Note that the best achievable depends on the Bayesian

[15]Let p_s^i be the posterior on the state being black given signal s, the expected accuracy of a guess is p_s^i $(1-p_s^i)$ if the guess is black (white). The optimality of the guess is coded as 0 if $p_s^i > 0.5$ and the guess is white or $p_s^i < 0.5$ and the guess is black, and 1 otherwise.

posterior associated with each signal realization. Specifically, the best achievable accuracy level is $\max\{p_s^i, 1-p_s^i\}$ (where p_s^i is posterior on the state being black given signal s from structure i). The figure shows that the accuracy of subjects' guesses on average closely tracks this best achievable level.[16,17] This suggests that subjects make near-perfect use of signals to inform their guesses. In particular, subjects make mistakes in guesses rarely and even when they do, they do so in cases where these mistakes are not very costly.

The high rate of guessing optimality is particularly clear in Figure 4b, which plots the optimality rate of guesses and depicts the distribution of this measure, the share of signals for which the subject's guess is optimal, computed at the individual level. The vast majority of subjects (77.3 percent) *always* make optimal choices. Calculating across subjects, 94.6 percent of the guesses are optimal. This near-perfect optimality strongly suggests that our methods for providing information typically avoid Bayesian errors (and other errors like confirmation bias), removing a major barrier to measuring information preferences.

Result 1 Subjects mostly make optimal use of information. Guesses conditional on the signal are optimal 94.6 percent of the time, and 77.3 percent of subjects make optimal guesses every time.

4.2 Impact of Instrumental Value on Information Demand

Next, we analyze the demand for information by examining ranking data. Since the specific set of seven structures varied randomly across subjects and ranking screens, and subjects were incentivized to focus on pairwise comparisons (Section 3.3), we use pairwise comparisons as the unit of analysis.[18,19] This approach allows us to systematically examine how differences in instrumental

[16] One outlier (as observed in Figure 4a) is a signal that consists of seven balls: six black and one white. Out of 55 subjects who randomly observed this signal (three of them twice), eight suboptimally guessed the color of the ball to be white. As discussed in Section 4.5, our main results are robust to removing such subjects from the analysis.

[17] Figure 11a in Appendix B plots the distribution of losses due to suboptimal guesses.

[18] While our primary analysis focuses on pairwise comparisons, we can also exploit random variation across ranking screens to examine how the full set of seven structures presented on each screen affects rankings. Table 6 in Appendix B provides such an analysis. Specifically, we use the number of structures that fully reveal the state (out of seven) as a measure of the simplicity of the ranking task to study whether this measure influences the main results reported in the paper. The analysis indicates that our main results remain robust when accounting for how this full-set-level measure moderates the impacts of instrumental value and other important structure characteristics (discussed in the next section) on demand for information.

[19] We also consider the full ranking of seven structures (rather than pairwise comparisons) as the unit of analysis. Table 7 in Appendix B presents the results of rank-ordered logit regressions that examine how the ranking of a structure within the set of seven is determined by its instrumental value and other important structure characteristics. These results are consistent with our analysis of pairwise comparisons.

value, and other structure characteristics, affect relative rankings.[20]

Figure 5a provides a first look at how demand for information is shaped by instrumental value. The figure examines all pairwise comparisons between information structures in which one structure is strictly more instrumentally valuable than the other. The bars show the likelihood with which the more instrumentally valuable structure was ranked higher than (as more preferred to) the other. We plot a separate bar for each possible value difference. When the value difference is relatively low, the optimal structure is preferred 67.2 percent of the time. This increases to 83.7 percent when the value difference is very high.

Figure 5b plots the cumulative distribution of optimality of pairwise comparisons at the individual level, focusing on cases where there is a strict value difference between the two information structures. In contrast to our above findings on guessing optimality (as seen in Figure 4b), there is substantial heterogeneity across subjects in ranking optimality. None of the 400 subjects in our data set achieves full optimality of 100 percent. While only 17.8 percent of subjects achieve an optimality rate above 90 percent, 46.8 percent of subjects achieve an optimality rate above 80 percent. It is also noteworthy that the optimality rate is below 60 percent for 27.3 percent of our subjects.[21]

Although these results show that (as expected) instrumental value is a major driver of information demand, they also reveal significant deviations from payoff-maximizing behavior. Overall, in 26.8 percent of relevant cases, subjects rank a strictly less instrumentally valuable information structure higher (i.e., as more preferred) than a more valuable information structure.[22,23] These mistakes come at significant costs to subjects. Conditional on a pairwise comparison being relevant for payment, ranking the less valuable structure above the more valuable one costs subjects between \$1 and \$4 in expected earnings, with an average loss of \$1.7.[24] These relatively severe errors in *comparing* information structures contrast sharply with the very low rate of errors in *using* information structures (to guess states) reported in the previous subsection.

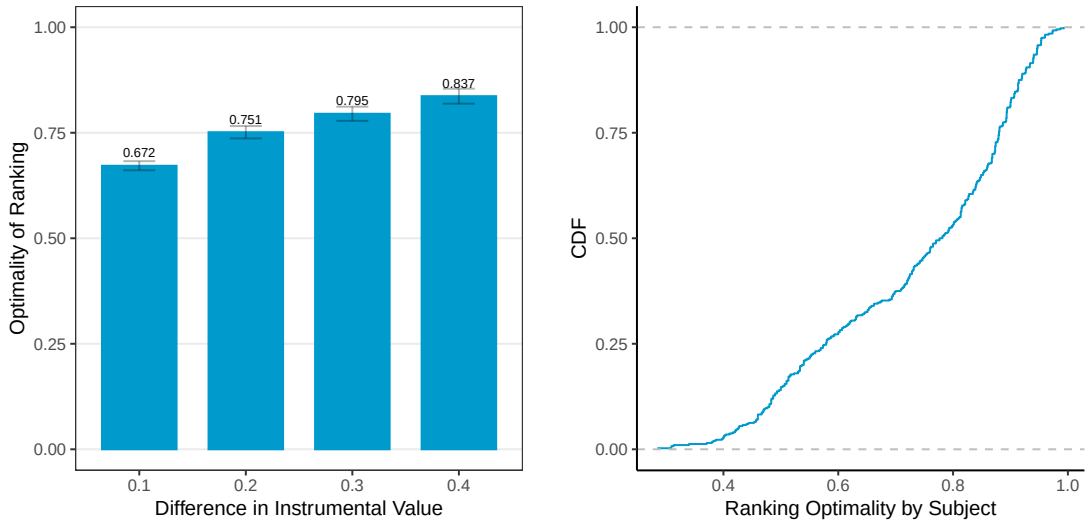
[20] Rankings within each screen impose transitivity over pairwise comparisons. To verify that this does not impact our main results, we replicate our analysis using a single random pairwise comparison from each ranking screen. While this approach substantially reduces the amount of data, it leaves our results unchanged.

[21] Figure 11b in Appendix B plots the distribution of losses due to suboptimal rankings.

[22] An error rate of this size is comparable to those observed in lottery choices: for instance, in Enke & Shubatt (2026), subjects fail to identify the lottery with the higher expected value in 27 percent of binary choices.

[23] There is no evidence of learning by introspection in the absence of feedback: the error rate remains stable across the eight ranking tasks, ranging from 26.0 to 27.8 percent. However, the amount of time subjects spend on each ranking screen declines with experience. This is not due to reduced attentiveness: the share of pairwise rankings that subjects change relative to the randomly generated original rank fluctuates only between 39.4 and 44.1 percent.

[24] To compute this, we assume optimal use of information by subjects.



(a) Optimality of Ranking by Value Difference (b) Distribution of Ranking Optimality

Figure 5: Optimality of Ranking Information Structures *Notes: We focus on pairwise comparisons. In panel (a) bars depict the rate at which one structure was preferred to another by value difference between the two structures. Gray lines denote 95 percent confidence intervals by bootstrapping. In panel (b) ranking optimality is computed at the individual level and denotes the share of pairwise comparisons (with strict value difference between two structures) for which ranking was optimal.*

Result 2 Ranking of information structures is strongly influenced by instrumental value, but there are frequent and costly deviations from payoff-maximizing behavior.

4.3 Structure Characteristics Beyond Instrumental Value That Impact Demand For Information

In this section, we ask which characteristics of information structures, beyond instrumental value, systematically predict information demand and what these relationships reveal about how subjects evaluate and compare information structures.

To motivate this question and set up our approach to answering it, consider a subject in our experiment tasked with comparing two information structures to determine which one to rank higher. When contrasting one structure with another, subjects might be behaviorally drawn to structures that are *simpler* relative to some criterion. For example, they might favor structures that are easier to evaluate, perhaps structures that contain fewer signals or generate fewer distinct posteriors. They might also approximate each structure's value using simple shortcuts rather than fully computing instrumental value and favor the structure that appears more valuable under this approximation. The key idea here is that differences between structures, beyond their instrumental value, may influence demand.

Alternatively, demand for information may not be driven by (or not solely by) *differences* between structures but by aggregate properties of the choice set. For instance, subjects might deviate more from optimal behavior when the overall *complexity* of the choice set increases, perhaps when the two structures are hard to compare, making it harder to determine which of the two is more valuable. A property of the choice set is symmetric in the two structures: it cannot favor either of them, but it can change how well the higher-value structure is identified. The key idea here is that aggregate-level characteristics of the choice set could serve as predictors of subjects' likelihood of making ranking errors.

We do not take an ex-ante stance on which characteristics are predictive of information demand. Instead, as outlined in Section 3, we constructed our dataset deliberately to introduce variation across a broad set of characteristics. Table 1 below provides a description of each characteristic varied by our design, and Table 3 in Appendix A presents the variation observed in our dataset with respect to each characteristic. The set covers primitive measures such as the number of signals and the probability of maximally certain or uncertain signals, information-theoretic measures such as entropy and skewness, and measures that may reflect potential simplification strategies, or non-standard preferences suggested by the literature. For each of eleven measures of a structure, we consider both the *difference* between two structures being compared and their *sum*, a property of the choice set. Together with two relational notions, Blackwell order and refinement (on partitions), and two measures of tradeoff complexity (which are properties of the choice set by construction) inspired by recent literature, excess dissimilarity and value-dissimilarity ratio, this gives 26 characteristics.

Impact of structure characteristics

We assess the predictive power of each structure characteristic for information demand in two complementary ways. First, we estimate how a one-standard-deviation increase in each characteristic, considered on its own, affects the optimality of rankings. Second, we use model *completeness* to assess how much of the explainable variation in rankings is captured by each characteristic.

For the completeness analysis, we start with a baseline logit model that uses the only characteristic theoretically relevant for rankings: which structure has higher instrumental value in a pairwise comparison. We then consider 26 alternative logit models, each of which augments the baseline with a single structure characteristic. Following Fudenberg, Kleinberg, Liang & Mullainathan (2022), completeness measures a model's reduction in prediction error relative to the baseline, scaled by the maximum achievable reduction in prediction error. We estimate this maximum achievable reduction using a flexible convolutional neural network (CNN).[25] A logit model that includes all

[25]Prediction error is measured by log loss: if p_{ij} is the predicted probability that structure i is ranked above j, the

Table 1: Characteristics of Information Structures

Structure characteristic	Description
Instrumental value (V)	The expected increase in guessing accuracy made possible by the decision maker being able to condition their action on the realized signals from an information structure.
Number of signals (n_s)	The number of distinct signals that an information structure can generate.
Number of signals with residual uncertainty (n_{ru})	The number of signals inducing non-degenerate posterior beliefs (i.e., beliefs other than 1 or 0) that an information structure can generate.
Number of distinct posteriors (n_{dp})	The number of distinct posterior beliefs that an information structure can generate.
Unweighted guessing accuracy ($\tilde{V}$)	The incorrect measure of expected increase in guessing accuracy that an information structure can produce when assuming the decision maker considers all possible signals realizing with the same chance.
Probability of maximally certain signals (P(max certain))	The probability that an information structure generates maximally certain signals (i.e., signals induce posterior of 1 or 0).
Probability of maximally uncertain signals (P(max uncertain))	The probability that an information structure generates maximally uncertain signals (i.e., signals induce posterior of 0.5).
Probability of worse-than-prior signals (P(worse than prior))	The probability that an information structure generates signals inducing posterior beliefs that are less certain (i.e., closer to uniform belief of 0.5) than prior beliefs.
Probability of guessing black (P(guess black))	The probability that an information structure generates signals inducing that the Bayesian optimal guess is 'black'—consistent with the Bayesian optimal guess if conditional only on the prior.
Skewness	The third normalized moment of Bayesian posterior $P(b\|s)$.
Informativeness (I)	The Shannon mutual information between prior beliefs and posterior beliefs induced by an information structure.
Blackwell order	The Blackwell order of information structures.
Refine	An information structure refines another if the former can produce the latter by arranging and combining signals (groups of the 10 balls) of the former in some way.
Excess dissimilarity	Each structure induces a compound lottery. We apply the measure to the first-stage lottery over signal-contingent guessing accuracies under optimal guesses. Excess dissimilarity is the CDF distance minus the absolute difference in expected values of these lotteries, adapting Enke & Shubatt (2026) to this setting.
Value-dissimilarity ratio	Using the same first-stage lotteries over guessing accuracies, the ratio of the absolute difference in expected values to the CDF distance, adapting Shubatt & Yang (2024) to this setting.

Notes: For the first 11 measures, we consider both the difference in the corresponding measure and the sum of the measure for a pair of information structures.

26 characteristics jointly has a completeness of 82 percent, indicating that the characteristics we consider capture most of the explainable variation in pairwise rankings.

Figure 6 compares the predictive power of the 26 structure characteristics in terms of (i) their absolute one-SD impacts on ranking optimality and (ii) their model completeness. The two measures are closely aligned, and the figure can be read as a ranking of the characteristics by their individual predictive power. The figure reveals substantial differences in predictive power across characteristics, with several characteristics beyond instrumental value emerging as strong predictors of rankings.[26] These strong predictors point to three distinct patterns in how subjects' rankings

loss is $-\ln p_{ij}$ when i is ranked higher and $-\ln(1-p_{ij})$ otherwise. We focus on pairwise comparisons with a strict difference in instrumental value, so the baseline logit model is equivalent to an intercept-only model. The CNN uses the 26 structure characteristics together with the raw partitions of the 10 balls corresponding to the two structures, allowing it to learn additional predictive features that may not be captured by the 26 characteristics. We evaluate all models using five-fold cross-validation, estimating each model on 80 percent of the data and evaluating its predictions on the remaining 20 percent in each fold.

[26]Some of these predictive relationships have a straightforward interpretation: larger differences in instrumental

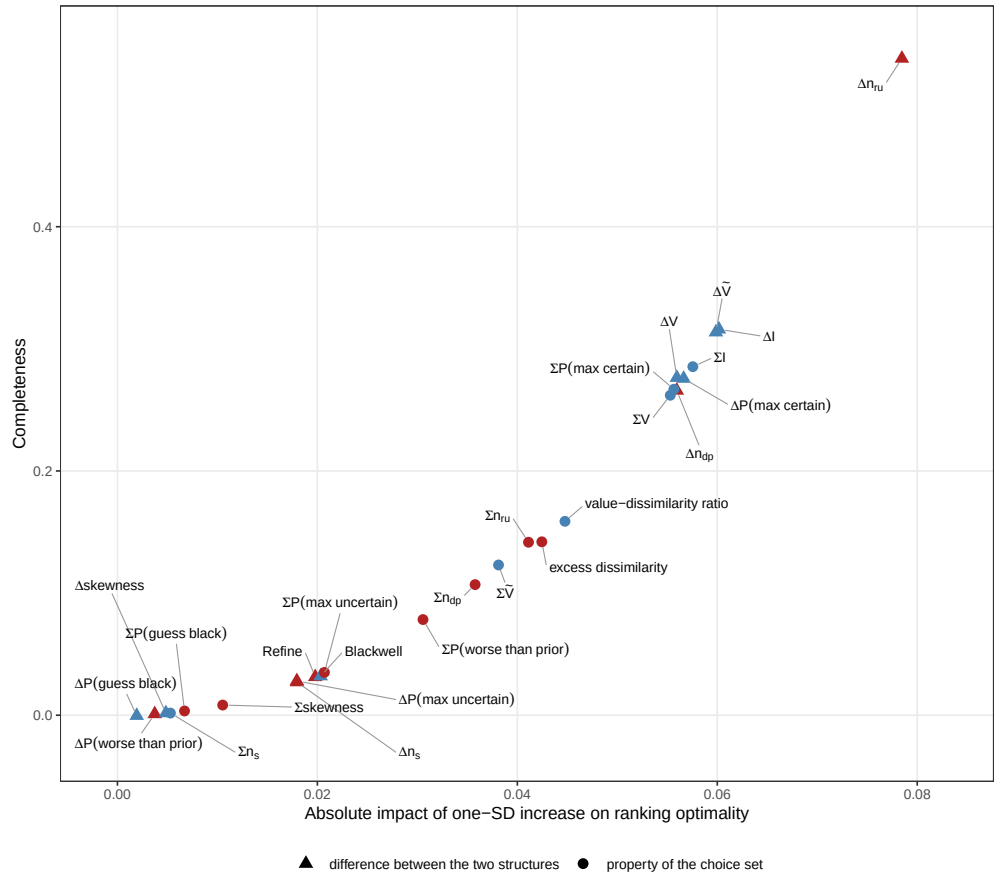


Figure 6: Impact and Completeness of Each Structure Characteristic *Notes: Each point represents one of the 26 structure characteristics. Triangles denote differences between structures and indicators of Blackwell dominance and refinement; circles denote properties of the choice set. Blue indicates a positive one-SD impact and red a negative impact. Both analyses use pairwise comparisons with a strict difference in instrumental value. Differences in characteristics are computed as the value for the higher-value structure minus that for the lower-value structure.*

depart from those implied by instrumental value and offer insight into how subjects evaluate and compare information structures.

First, differences in the *size* of the two information structures being compared are strongly predictive of rankings. One characteristic clearly stands out: Δn_{ru}, the difference in the number of signals with residual uncertainty, has the largest absolute one-SD impact of all 26 characteristics, at 7.8 percentage points, and a completeness of 54 percent on its own. Its impact is negative, as indicated by its red color: subjects are less likely to prefer a higher-value structure to a lower-value one when the former generates more signals with residual uncertainty. Differences in the number of distinct posteriors (Δn_{dp}) and signals overall (Δn_s) have impacts in the same direction, though

value (ΔV) predict higher ranking optimality, as do larger differences in informativeness (ΔI), which is strongly correlated with ΔV (Pearson's $r = 0.88$).

smaller in magnitude. Since assessing an information structure's value involves determining the distribution of actions it induces and their payoff consequences in terms of guessing accuracy, the size of the structure can plausibly affect the complexity (difficulty) of this evaluation. We therefore interpret the negative impacts of these size measures as indicating a form of *complexity aversion*.

Second, one of the most predictive characteristics suggests that subjects may be simplifying the assessment of instrumental value by neglecting signal likelihoods. The difference in unweighted guessing accuracy ($\Delta\tilde{V}$) has a positive one-SD impact of 6.0 percentage points and a completeness of 31 percent on its own. This characteristic captures the difference in the accuracy induced by the two structures that would be inferred if the subject were to heuristically ignore the likelihood of signals. Specifically, unweighted guessing accuracy quantifies the value each structure would have if decision makers correctly computed the posterior for each signal but assumed that all signals were equally likely rather than weighting them by their true likelihood. This is a natural way to simplify the problem, easing the challenge of weighting and aggregating guessing accuracies across signals. As we show below, $\Delta\tilde{V}$ remains predictive of rankings even after controlling for the difference in instrumental value. We interpret this pattern as indicating a form of *heuristic simplification* in how subjects evaluate information structures.

Third, properties of the choice set as a whole predict how strongly subjects respond to differences in instrumental value. Higher total instrumental value ($\sum V$) and informativeness ($\sum I$), as well as lower tradeoff complexity (lower excess dissimilarity or a higher value-dissimilarity ratio), predict higher ranking optimality. Thus, these relationships reflect *attenuation*: subjects have more difficulty identifying the higher-value information structure when the overall value of the choice set is low or the structures are harder to compare.

Together, these findings document three distinct responses to complexity—aversion, simplification and attenuation—within a unified experimental framework. Our results suggest that each plays an important role in how individuals evaluate and compare information structures.

Representative characteristics

To illustrate the patterns discussed above more concretely, we focus on four representative characteristics that are strongly predictive of rankings and easy to interpret. We use the differences in the number of signals with residual uncertainty (Δn_{ru}) and distinct posteriors (Δn_{dp}) to illustrate complexity aversion, the difference in unweighted guessing accuracy ($\Delta\tilde{V}$) for heuristic simplification, and total instrumental value ($\sum V$) for attenuation. Together, these four characteristics achieve a completeness of 75 percent, or 92 percent of the completeness achieved by all 26

characteristics.[27,28] Given the correlations among characteristics, these four should be viewed as representative predictors rather than a uniquely important set.

Because these characteristics can be correlated with differences in instrumental value, Figure 7 examines how each predicts rankings when the value difference is held fixed. Each panel plots the rate at which one structure is preferred to another as a function of their value difference, with separate curves for comparisons in which the relevant characteristic is strictly above its sample median or weakly below it. Gaps between the curves at a given value difference therefore show how rankings vary with the characteristic beyond instrumental value.[29] Under optimal behavior, both curves would follow the same step function, rising from zero to one at zero.

Panels (a) and (b) show that, holding the difference in instrumental value constant, subjects are more likely to prefer structures that induce fewer signals with residual uncertainty and fewer distinct posteriors: the curve shifts downward at every value difference when a structure has more of these elements than the other. These effects are substantial and particularly pronounced when the difference in instrumental value is small. Among pairwise comparisons with a strict difference in instrumental value, subjects prefer the higher-value structure to the lower-value one 74 percent of the time when the former induces weakly fewer signals with residual uncertainty, but only 48 percent when it induces strictly more. The corresponding rates for the number of distinct posteriors are 76 percent versus 65 percent.

Panel (c) illustrates the impact of unweighted guessing accuracy. Among pairwise comparisons with a strict difference in instrumental value, subjects prefer the higher-value structure to the lower-value one 75 percent of the time when the former has strictly higher unweighted guessing accuracy, but only 46 percent when it has weakly lower unweighted guessing accuracy. Although unweighted guessing accuracy and instrumental value are closely related by construction, panel (c) shows that rankings vary substantially with $\Delta\tilde{V}$ even when ΔV is held fixed.

[27] LASSO and decision-tree analyses support the choice of these representatives (Appendix C): the first four characteristics to enter the logistic LASSO are Δn_{ru}, $\sum I$, $\Delta\tilde{V}$ and Δn_{dp}; the fifth adds less than one percentage point of completeness. These characteristics also appear near the top of classification decision trees. Given their high correlation ($r = 0.96$), we use $\sum V$ in place of $\sum I$ for its more direct interpretation in our setting. Using $\sum I$ in panel (d) of Figure 7 produces nearly identical results (Appendix B, Figure 13).

[28] Table 5 in Appendix B reports logit regressions of ranking optimality on each of the four characteristics and on all four together.

[29] Here we use graphs to show how rankings vary with each characteristic at a fixed difference in instrumental value. Regressions controlling for the value difference yield similar findings. Controlling for ΔV, the one-SD impacts of Δn_{ru}, Δn_{dp}, $\Delta\tilde{V}$, and $\sum V$ in separate regressions are -6.9, -5.3, 4.1, and 5.5 percentage points, and the corresponding impacts of ΔV are 1.6, 5.3, 2.5, and 5.6.

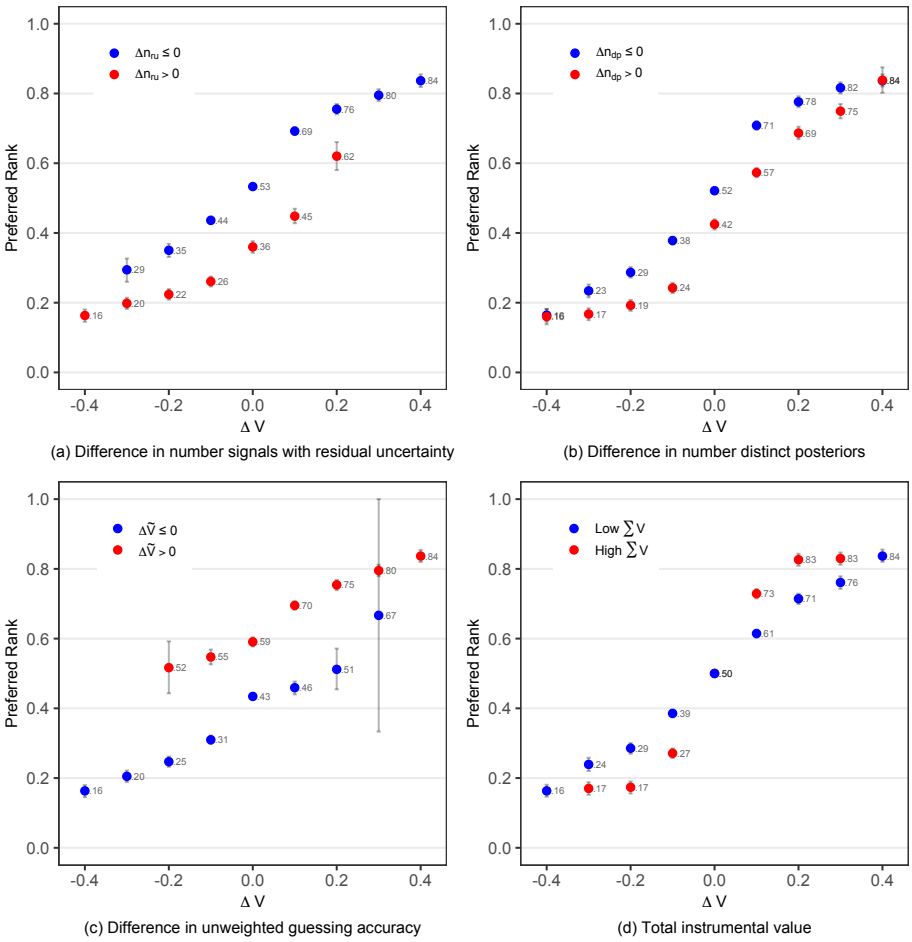


Figure 7: Pairwise Ranking of Information Structures by Value Difference *Notes: We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

Finally, panel (d) shows the impact of the total instrumental value of the two structures, $\sum V$, on ranking optimality. Since this characteristic operates at the aggregate level, reflecting properties of the choice set, its influence produces a distinct effect, visible in the figure. Specifically, $\sum V$ predicts how closely the response curve resembles the theoretically predicted step function: at lower values of $\sum V$, the curve is flatter, indicating attenuation. Among pairwise comparisons with a strict difference in instrumental value, subjects prefer the higher-value structure to the lower-value one 70 percent of the time when $\sum V$ is low, compared with 78 percent when it is high.[30] In

[30]We examine two potential alternative drivers of this result. (1) Subjects may rank their top choices more carefully than their bottom choices. To test this, we compare the shares of pairwise reversals relative to the initially presented ranking among the top and bottom three structures in each subjects final ranking. Consistent with this concern, we

other words, subjects' responsiveness to differences in instrumental value increases with the total instrumental value of the choice set, suggesting that comparisons between high-value structures are cognitively easier. This is intuitive: high-value structures are typically associated with signals that fully reveal the state, whose contribution to value is immediate, while low-value structures tend to generate ambiguous signals, whose contributions must be computed and aggregated ($\sum V$ has a correlation of -0.78 with $\sum n_{ru}$ in our sample), making comparisons between information structures more difficult. A similar attenuation pattern emerges for the two measures of tradeoff complexity (i.e., the difficulty of trading off competing considerations), as shown in Figure 12 of Appendix B (Enke & Shubatt 2026, Shubatt & Yang 2024).

Explanatory power of the representative characteristics at the individual level

The evidence so far shows that structure characteristics beyond instrumental value predict rankings at the aggregate level. We next ask whether simple decision rules based on these characteristics fit individual subjects' rankings better than the theoretical benchmark of always choosing the structure with higher instrumental value. We consider three alternative rules: choose the structure with (i) fewer signals with residual uncertainty, (ii) fewer distinct posteriors, or (iii) higher unweighted guessing accuracy, rather than the structure with highest instrumental value. These rules correspond to the first three representative characteristics in Figure 7.[31]

For each alternative rule, we restrict attention to pairwise comparisons in which the two rules strictly disagree: one structure has higher instrumental value, but the alternative rule strictly favors the other. We classify a subject as more consistent with the alternative rule if they follow it in more than half of these comparisons, more consistent with the optimal rule if they follow it in fewer than half, and equally consistent with both if they follow each rule equally often.[32]

Each row of Table 2 compares one alternative rule with the optimal rule, showing the number of pairwise comparisons per subject in which they strictly disagree, the share of subjects more consistent with each rule, and how consistently subjects in each group follow the better-fitting rule. Although relatively few comparisons distinguish the residual-uncertainty and unweighted-accuracy

find more reversals on average among the top three. However, a substantial share of subjects (32.8 percent) show the opposite pattern—making (weakly) more changes among the bottom three—and our main results replicate within this subset. (2) As documented in the previous section, subjects make fewer mistakes when comparing maximally valuable information structures that induce guessing accuracy of 100 percent to others. Removing such comparisons from the analysis does not eliminate the predictive power of total instrumental value either.

[31] The last characteristic, $\sum V$, is a property of the choice set and does not by itself favor either structure.

[32] The equal-fit group includes subjects with no pairwise comparisons in which the alternative and optimal rules strictly disagree, who account for at most 1 percent of subjects for each rule.

Table 2: Individual-Level Consistency with Alternative and Optimal Rules

	Comparisons where the rules disagree	Share of subjects (%)			Consistency (%)	
Alternative rule	Mean (median)	Alternative fits better	Optimal fits better	Equal fit	Alternative group	Optimal group
Fewer signals with residual uncertainty	6.6 (6)	46	42	13	77	74
Fewer distinct posteriors	39.8 (39)	23	75	2	62	74
Higher unweighted guessing accuracy	5.2 (5)	50	35	16	79	75

Notes: The analysis uses pairwise comparisons in which the alternative and optimal rules strictly favor different structures. Consistency is the share of these comparisons in which a subject follows the better-fitting rule, averaged equally across subjects within each group.

rules from optimal choice, the median subject still faces six and five such comparisons, respectively. 46 and 50 percent of subjects are more consistent with the alternative rules than with the optimal rule, respectively, with average consistency rates of 77 and 79 percent within these groups. The distinct-posterior rule disagrees with optimal choice more frequently and fits a smaller but still substantial share of subjects better (23 percent), with an average consistency rate of 62 percent within this group. These findings show that the aggregate patterns also appear in individual rankings: for substantial shares of subjects, simple rules based on structure characteristics fit choices better than the instrumental-value benchmark when the rules disagree.

We summarize the results from this section below:

Result 3 Structure characteristics beyond instrumental value are predictive of preference rankings of information structures. Subjects display aversion to larger structures, such as those that produce more distinct posteriors or more signals with residual uncertainty, appear to simplify valuation by neglecting signal likelihoods, and respond less to value differences when comparisons are difficult.

4.4 Impact of Structure Characteristics Relative to Participant Characteristics

To help contextualize the preceding results, we next examine characteristics of *decision makers* that predict ranking behavior. We consider 16 participant characteristics measured in the experiment (see Section 3), which are listed in Table 4 of Appendix A. They cover cognitive ability, measured by the overall score on the ten cognitive questions, and attitudes towards risk, ambiguity, and compound lotteries. They also include subjects computations of the instrumental value of an information structure and the winning chance of a compound lottery, along with the mistakes made in these computations. Finally, they capture subjects reported confidence in their rankings

and in both computations.

To assess the predictive power of these participant characteristics and compare it with that of the structure characteristics, we follow the same procedure as in the previous section. We focus on all pairwise comparisons where one structure is of strictly higher value than the other and consider each of the 16 participant characteristics separately. For each characteristic, we estimate the impact of a one-standard-deviation increase on ranking optimality and the completeness of a model that adds the characteristic to a baseline logit model using only an indicator for which structure has higher instrumental value.[33]

Figure 8 plots the two measures. Two of the participant characteristics stand out: (i) cognitive score, which can be interpreted as a measure of intelligence (or alternatively as a measure of engagement with the experiment), has a one-SD impact of 7.8 percentage points on ranking optimality and a completeness of 23 percent on its own;[34] (ii) confidence in ranking optimality, which measures how certain subjects are in the optimality of their rankings, follows with a one-SD impact of 5.6 percentage points and a completeness of 11 percent.[35] Subjects with higher cognitive scores and those who are more confident in their rankings are more likely to rank the higher-value structure as more preferred.

How do these impacts compare with those of the structure characteristics? The strongest characteristic on each side has the same impact size: a one-standard-deviation increase in the cognitive score raises ranking optimality by 7.8 percentage points, and a one-standard-deviation increase in Δn_{ru} lowers it by 7.8 percentage points. The other three representative structure characteristics, $\Delta\tilde{V}$, Δn_{dp}, and $\sum V$, have absolute one-SD impacts of 6.0, 5.6, and 5.5 percentage

[33]Completeness is measured against different benchmarks for participant and structure characteristics. For participant characteristics, the benchmark is a flexible neural network using the 16 participant characteristics, trained under the same procedure as the CNN in Section 4.3. These characteristics distinguish subjects, allowing the network to capture differences in their average ranking optimality. The benchmark of Section 4.3, by comparison, uses only information about the structures. The two benchmarks therefore imply different achievable reductions in prediction error, so completeness values cannot be compared directly across the two sets of characteristics. We instead compare their one-SD impacts on ranking optimality. Figure 16 in Appendix B provides a completeness comparison using a common benchmark: a CNN trained on the raw partitions, the 26 structure characteristics, and the 16 participant characteristics together.

[34]The cognitive score sums the scores on the three Raven's test questions, the three belief bias syllogism questions, the three cognitive reflection test questions, and the Wason selection task. On their own, these four components have one-SD impacts of 8.2, 5.2, 3.5, and 0.8 percentage points, respectively.

[35]The next two characteristics, the absolute mistakes in the computation of the winning chance of the compound lottery and in the computation of the instrumental value of an information structure, have one-SD impacts of -3.3 and -3.1 percentage points. Other participant characteristics have an absolute impact below 3 percentage points.

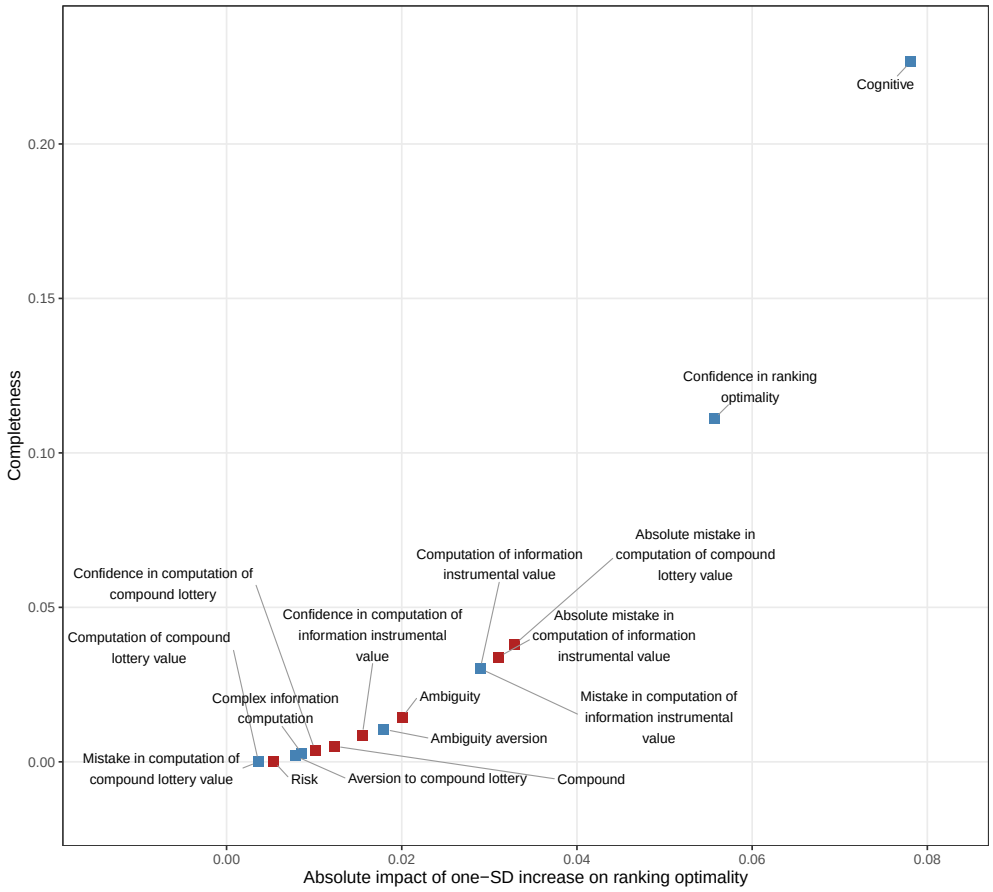


Figure 8: Impact and Completeness of Each Participant Characteristic *Notes: Each point is one of the 16 participant characteristics listed in Table 4. The horizontal axis is the absolute value of the impact of a one-standard-deviation increase in the characteristic alone on ranking optimality. The vertical axis is the completeness of a model that adds the characteristic alone to the baseline, where the irreducible error is given by a neural network whose inputs are the 16 participant characteristics (see footnote 33). Blue color denotes a positive one-SD impact and red a negative impact. Both analyses focus on all pairwise comparisons with a strict value difference.*

points, each nearly the same as or even larger than the impact of confidence in ranking optimality and larger than the impact of any other participant characteristic.[36] These findings highlight both the role of individual differences – some participants are better able and more willing to identify higher-value structures – and the persistent influence of structure characteristics despite these differences.

[36]Figure 16 in Appendix B, which measures both kinds of characteristics against one common benchmark, tells the same story in terms of completeness. Δn_{ru} and the cognitive score lead with a completeness of 13 percent each, followed by a group of structure characteristics (including the other three representative structure characteristics) with a completeness between 7 and 8 percent, and then by confidence in ranking optimality with 6 percent.

Interactions between structure characteristics and participant characteristics

The impacts of structure characteristics reported so far are averages across subjects who differ in their propensity to rank structures optimally. A natural question is whether structure characteristics are more predictive of rankings for some types of subjects than others. Answering this question is challenging, as participant characteristics and structure characteristics can interact in many different ways to jointly influence ranking optimality. One intuitive expectation is that structure characteristics beyond instrumental value matter mainly for less sophisticated subjects. Those with lower cognitive scores or lower confidence make more mistakes, and the structure characteristics identified in the previous section could systematically predict when these mistakes occur. In contrast, these characteristics may be less predictive for more sophisticated subjects, who make fewer mistakes. We show in this section that, to the degree that such interactions are present, they do not arise in this anticipated way.

To examine whether structure characteristics predict rankings more strongly for some subject types than others, we split subjects into two groups at the median of a participant characteristic and estimate the one-SD impact of each of the four representative structure characteristics on ranking optimality separately in the two groups.[37] We first focus on pairwise comparisons in which the difference in instrumental value between the two structures is 0.1, the smallest strictly positive difference in our design. These comparisons account for 42 percent of all pairwise comparisons with a strict value difference and have the highest frequency of deviations from optimal behavior. In each, selecting the suboptimal structure reduces guessing accuracy by 10 percentage points, allowing us to compare the impacts of structure characteristics across the two groups while holding the cost of a mistake fixed.

The results of this analysis are presented in Figure 9, with subjects split by their cognitive score in the top panel and by their confidence in ranking optimality in the bottom panel. Two key observations emerge. First, all four representative structure characteristics have a sizable and significant impact in both groups, whether subjects are split by cognitive score or by confidence. Second, their impacts are larger, not smaller, for subjects with high cognitive scores or those with high confidence in ranking optimality. When we extend the analysis to all pairwise comparisons with a strict value difference (Figure 17 in Appendix B), the differences between the two groups become smaller, as expected: at large value differences, ranking optimality is close to the ceiling for most subjects, particularly the more sophisticated ones. Nonetheless, the differences point in

[37] For each structure characteristic, we regress ranking optimality on the standardized structure characteristic, an indicator for whether the subjects participant characteristic is strictly above the median, and their interaction.

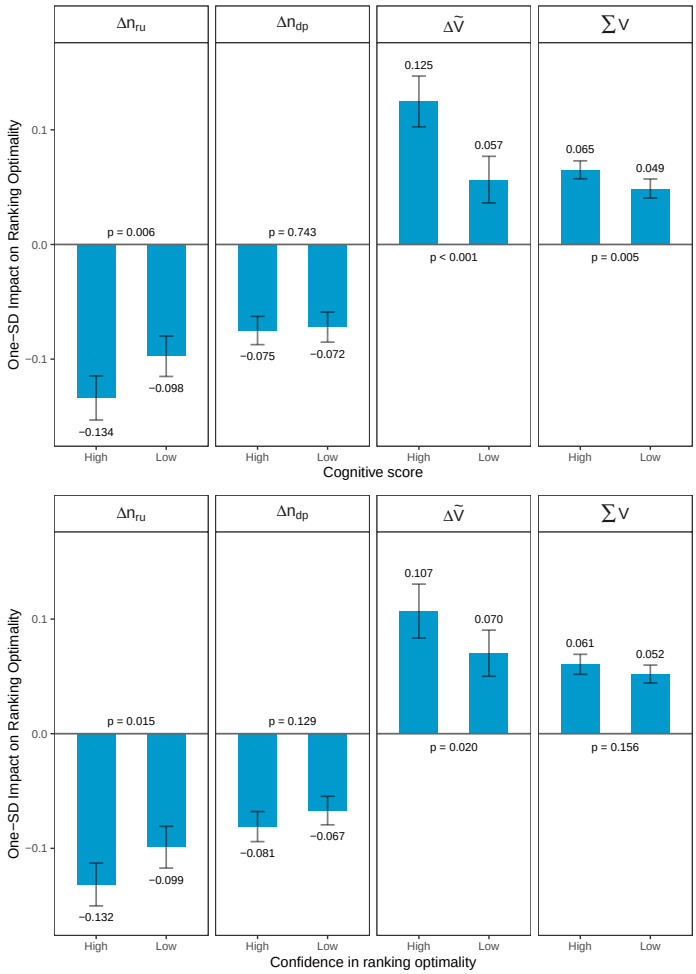


Figure 9: Impact of Structure Characteristics by Subject Type *Notes: Analysis is on pairwise rankings where* $\Delta V = 0.1$. *Each bar is the impact of a one-standard-deviation increase in the structure characteristic on ranking optimality, estimated separately for subjects whose cognitive score (top panel) or confidence in ranking optimality (bottom panel) is strictly above the median value in our sample (High) and for those for whom it is weakly below (Low). Gray lines denote 95 percent confidence intervals, with standard errors clustered by subject. p-values are for the test that the impacts in the two groups are equal.*

the same direction in six of the eight by-group comparisons, the two exceptions being small and insignificant differences for Δn_{dp}.

Overall, we find that subjects who are generally more aligned with optimal rankings—due to intelligence, engagement, or confidence—are also the ones whose behavior is most shaped by structure characteristics that should, in principle, be irrelevant to optimality. This suggests that deviations from optimal rankings are not simply driven by lack of ability or effort but may instead reflect systematic tendencies in how even high-performing individuals process the value of different information structures. This may seem surprising: subjects who rank structures more optimally

on average are also more influenced by structure characteristics that are irrelevant to optimality.[38] However, as we discuss in Section 5, this pattern has a fairly natural (and interesting) interpretation. What the results suggest is that mistakes by less sophisticated subjects tend to be noisier and more random (and therefore less explained by any deterministic metric), while mistakes by more sophisticated subjects tend to be more systematic and predictable.

Finally, we examine the role of ambiguity aversion. Although ambiguity aversion does not stand out among the participant characteristics in Figure 8, with a one-SD impact of 1.8 percentage points on ranking optimality, it allows us to connect our findings to recent work on ambiguity and complexity. When we separate subjects by their ambiguity attitudes, a similar pattern emerges: ambiguity-averse subjects tend to rank structures more optimally, yet their rankings are more strongly predicted by the four structure characteristics than those of ambiguity-neutral or ambiguity-loving subjects (Figure 18 in Appendix B). This may point to a role for ambiguity attitudes in driving complexity aversion in the evaluation of information structures (see de Clippel et al. (2024) for recent evidence on the link between ambiguity attitudes and complexity aversion).

Result 4 Structure characteristics are (at least) similarly predictive of rankings as participant characteristics that could influence demand for information. Moreover, their impacts are similarly large, and in many cases larger, for subject subgroups that are more likely to exhibit optimal behavior.

4.5 Robustness of Findings and Additional Experiments

Robustness across subsamples

The findings in the previous section indicate that structure characteristics exhibit stronger predictive power over rankings for certain subpopulations. To ensure that our main results are not driven by inattentiveness, lack of engagement, or general confusion with the experiment, we replicate Figure 7 for different subsamples in Appendix B. Figure 14 restricts the analysis to subjects who answered all comprehension questions correctly on their first attempt, and Figure 15 focuses on subjects who made perfectly optimal guesses (across all 15 questions) in the guessing portion of the experiment. The patterns in these subsamples closely align with those observed in the full sample, reinforcing the robustness of our main findings.

[38] Figure 19 in Appendix B studies more closely the estimated impact of these selected structure characteristics as a function of the overall optimality rate in rankings. The non-monotonic relationship depicted in the figure is reminiscent of similar findings in Gonçalves (2024).

Connections to guessing behavior

While results in Section 4.1 indicate that subjects rarely make mistakes in guessing questions, such mistakes do occasionally occur. We examine the extent to which structure characteristics predict these errors, as well as their impact on response times, that is, the time it takes subjects to determine their guess.[39] Results from this analysis are reported in Figure 20 of Appendix B. Overall, we find guesses are more likely to be optimal for information structures with high instrumental value, fewer signals with residual uncertainty, and higher unweighted guessing accuracy. The impact of these variables on response times follows the opposite pattern: subjects respond more quickly when structures have high instrumental value, fewer signals with residual uncertainty, and high unweighted guessing accuracy.[40,41] Overall, we find that the structure characteristics that influence ranking behavior appear to be linked to those that make it more challenging for subjects to determine optimal guesses given a signal.

Laboratory experiment

An earlier version of this paper reported results from a laboratory experiment with a similar design (further details in Appendix D). A key difference is that all subjects in the lab study saw the same 16 information structures, chosen to vary certain structure characteristics (e.g., number of signals, entropy informativeness) while holding instrumental value constant. This limited the variation in structure characteristics. The current study addresses this by expanding the set of information structures from 16 to 237.

Behavior in the laboratory study closely mirrors that of the main study.[42] Most importantly, as shown in Figure 22 of Appendix D, the four representative structure characteristics in the main study similarly predict ranking behavior in the lab data. These findings reinforce the robustness of the main study's results, confirming that they are unlikely to be driven by specific details of the

[39] Since we analyze guesses conditional on signal realization, we can also directly assess whether characteristics of the signal itself, rather than just structure-level characteristics, influence guessing accuracy and response times.

[40] We do not find the number of distinct posteriors associated with an information structure to significantly affect guessing accuracy or response time.

[41] Similar patterns emerge with participant characteristics. Guessing optimality is higher among subjects with higher cognitive scores and, to a lesser extent, among those who are more confident in their rankings, and subjects with higher cognitive scores also tend to respond more quickly. In fact, as shown in Figure 20 in Appendix B, the cognitive score has a much greater impact on guessing optimality than any structure-level or signal-level variable.

[42] For instance, subjects guess optimally 98.8 percent of the time and rank less valuable structures above more valuable ones in 22.8 percent of relevant pairwise comparisons. These rates closely align with those from the main study (94.6 percent and 26.8 percent) and with those from a dedicated replication study using the same interface and subject pool as the main study but limited to the 16 structures used in the lab study (93.2 percent and 25.9 percent).

experimental interface or the characteristics of the online subject pool.[43]

In addition to gathering *ranking* data (as in our main study), in the lab study we also asked subjects to explicitly *value* information structures via elicited willingness to pay (WTP) measures. WTP responses differ from ranking results in a few ways. They ordinally violate instrumental value in 39.7 percent of relevant pairwise comparisons, exceeding the 22.8 percent violation rate observed in the ranking data. Furthermore, the pairwise orderings implied by WTP agree with the rankings in only 74.9 percent of comparisons. Moreover, the impacts of the four structure characteristics discussed above on WTP do not always align with their impacts on rankings (see Figures 27 and 28 in Appendix D). These findings are consistent with evidence that choice tasks, such as our ranking task, can be less noisy than explicit valuation tasks (see, e.g., Bouchouicha et al. (2025)) and with a long literature, particularly in risky choice, suggesting that the two types of tasks can involve different cognitive processes and yield different implied preferences (see, e.g., Lichtenstein & Slovic (1971), Grether & Plott (1979)).[44]

We also ran two treatment variations in the laboratory study that are useful in understanding the mechanism behind our main results. First, we examine to what extent difficulties in evaluating information structures are due to the difficulty of forecasting how those information structures will be used to improve behavior (see Niederle & Vespa (2023) for a review of the literature on failures in contingent reasoning like these). Evaluating an information structure requires the decision maker to anticipate, for each signal it may generate, which guess the signal will induce, and to aggregate the resulting guessing accuracies over signals. If subjects' difficulties lie in the first step, i.e., in anticipating their own use of the signals, then having already determined the guess for each signal should make the evaluation easier. To test this, in the laboratory study we included a Reverse treatment in which subjects made guesses for all possible signal realizations before ranking the

[43] In an earlier version of this paper that was based on the laboratory data, we found that differences in entropy informativeness were an important predictor of pairwise rankings. However, this finding turns out to be special to the restricted set of 16 information structures selected for that earlier study (indeed, we find the same result in a replication study (described in Appendix D) that uses the 16 structures from the laboratory study but uses the Prolific subject pool and the experiment interface of the main study). When we dramatically expand the set of information structures to 237, we find instead that Δn_{ru}, Δn_{dp}, and $\Delta \tilde{V}$, rather than differences in entropy informativeness, are better and more robust predictors of behavior.

[44] We elected not to include WTP elicitations in our main study for several reasons. First, ranking data is less noisy (and closer to optimal behavior) and we decided it would therefore be a better fit for an online sample. Second, we think ranking data is more similar to the types of choices over information structures people typically make in the field (people are rarely asked to attach explicit monetary values to information). Finally, in comparison to the lab study, we wanted to study preferences over far more information structures, which is relatively easy with ranking tasks but would be difficult using WTP elicitations.

structures, and were later reminded of those guesses during the ranking task, so that the first step of the evaluation was completed and displayed to them. Under the hypothesis that our main findings are driven by subjects' failure to contingently reason about how different signals induce different guesses, we would expect this treatment to eliminate or at least reduce the severity of our results. However, the Reverse treatment has no effect on ranking errors: the same structure characteristics continue to strongly influence rankings (see Figures 24 and 25 in Appendix D). The difficulties we document thus lie in the aggregation of the guesses over signals, not in the anticipation of the guesses themselves, which is consistent with the near-optimal guessing behavior documented in Section 4.1.

Second, we examine whether non-standard preferences over risk, loss, or the timing of uncertainty resolution could explain the systematic deviations from optimal behavior observed in the ranking of information structures. For instance, a structure that generates many signals with residual uncertainty leaves the decision maker exposed to the risk of a wrong guess after the signal is observed, and an aversion to such structures could reflect anticipated regret or a preference about when uncertainty is resolved rather than the difficulty of evaluating them. To test this, the laboratory study included a *No Uncertainty* treatment, which eliminated uncertainty using methods similar to those used in Martinez-Marquina, Niederle & Vespa (2019) and Oprea (2024) while preserving the cognitive demands of the task. Instead of drawing one ball, we drew all ten and paid subjects for each correct guess, making the payoff from an information structure a deterministic amount equal to the guessing accuracy it induces and eliminating risk, regret, and the resolution of uncertainty from its use. Despite these differences, ranking patterns remain similar in this treatment (see Figure 26 in Appendix D, which replicates Figure 7). This suggests that preferences over timing or risk are unlikely to be the primary drivers of deviations from optimal ranking behavior in our data. We note that the treatment removes the objective uncertainty of the task, not the uncertainty a subject may perceive about the value of a structure while evaluating it. Such internal uncertainty may help explain the attenuation pattern documented in Section 4.3: when structures are harder to evaluate, subjects perceptions of the difference in their values may be noisier, making rankings less responsive to the true value difference.

Alternative information representation and preference elicitation

Finally, we examine whether the patterns we identify depend on the partition representation of information structures or on the use of ranking tasks to elicit preferences. To do so, we conduct an additional experiment with 200 Prolific subjects, presenting information structures in a more standard tabular format and eliciting preferences through binary choices.

Figure 1(b) provides an example of how information structures were presented in this experiment. Each row represents a possible signal realization and numerically specifies its likelihood and the probabilities of a black and a white ball conditional on that signal. Following the structure of the main experiment, subjects first complete guessing tasks and then continue to make 80 binary choices between pairs of structures drawn from the same universe of 237 information structures.[45]

The alternative design does not improve subjects' ability to use or choose information. Guesses remain highly optimal, at 90.1 percent, but this rate is somewhat lower than the 94.6 percent observed in the main study. Subjects choose the less valuable structure in 32.3 percent of binary choices with a strict value difference, compared with 26.8 percent of pairwise rankings in the main study. Choices are also less responsive to differences in instrumental value: the likelihood of choosing the higher-value structure rises from 65 percent at $\Delta V = 0.1$ to 71 percent at $\Delta V = 0.4$, compared with 67 to 84 percent in the main study (Figure 29 in Appendix E). These results suggest that the partition representation and ranking tasks used in the main study do not hinder—and may even facilitate—optimal behavior.

We next examine whether structure characteristics beyond instrumental value continue to predict information demand under the alternative design. Figure 10 repeats the analysis of Figure 6 using data from this experiment: for each of the 26 structure characteristics, we estimate its one-SD impact on choice optimality and its completeness.[46] A model with all 26 characteristics has a completeness of 97 percent, suggesting that the characteristics we consider capture most of the explainable variation under the alternative design as well.

More importantly, under the alternative design, subjects continue to display aversion to larger structures, appear to simplify valuation by neglecting signal likelihoods, and respond less to value differences when comparisons are difficult. These three patterns documented in Section 4.3 persist even though the magnitudes and relative importance of individual structure characteristics differ from those in the main study.

First, subjects continue to favor smaller information structures. As in the main study, Δn_{ru} has the largest one-SD impact (-6.6 percentage points) and completeness (53 percent). Among choices between structures with a strict value difference, subjects choose the higher-value structure 69 percent of the time when it induces weakly fewer signals with residual uncertainty, but only 45 percent when it induces strictly more (compared with 74 and 48 percent in the main study). Two

[45] The pairs were drawn in a way to facilitate direct comparison with the main experiment (see Appendix E for details).

[46] Completeness is estimated as described in Section 4.3, using a CNN trained on data from the new experiment as the benchmark.

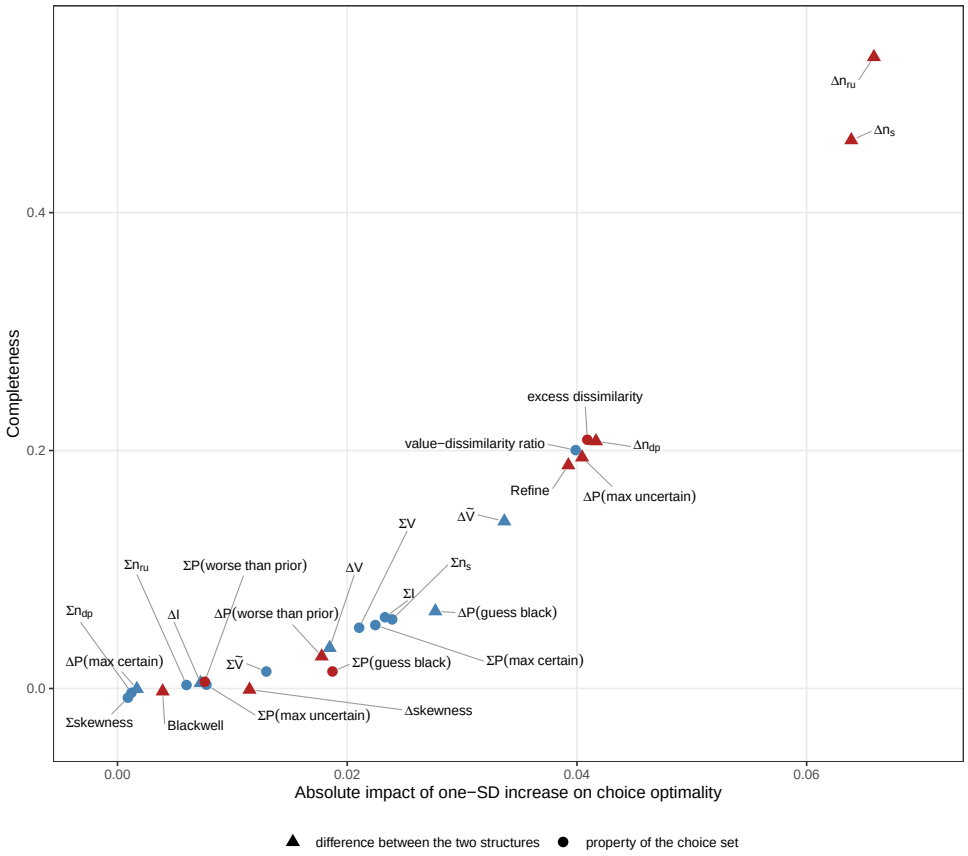


Figure 10: Impact and Completeness of Each Structure Characteristic under the Alternative Design *Notes: This figure repeats the analysis in Figure 6 using data from the alternative design. Axes, symbols, and colors follow the conventions in Figure 6.*

other measures of differences in structure size, Δn_s and Δn_{dp}, also have sizable negative impacts on choice optimality and are among the most predictive characteristics. The number of signals plays a more prominent role than in the main study, possibly because the tabular representation makes it harder to recognize signals that induce identical posteriors.

Second, subjects continue to favor structures with higher unweighted guessing accuracy, even though signal likelihoods are explicitly reported in the table. The difference in unweighted guessing accuracy, $\Delta\tilde{V}$, has a one-SD impact of 3.4 percentage points and a completeness of 14 percent. Among choices with a strict value difference, subjects choose the higher-value structure 69 percent of the time when it also has strictly higher unweighted guessing accuracy, but only 51 percent otherwise (compared with 75 and 46 percent in the main study). These findings remain consistent with subjects simplifying the choice task by neglecting signal likelihoods, even when those likelihoods are explicitly reported.[47]

[47]The difference in the probability of maximally uncertain signals, ΔP(max uncertain), has a relatively strong

Third, subjects continue to respond less to value differences when comparisons are difficult, although the characteristics that best capture this attenuation differ across designs. Whereas total instrumental value was a prominent predictor in the main study, the two tradeoff complexity measures play a more prominent role under the alternative design. Excess dissimilarity and value-dissimilarity ratio have the largest impacts among the properties of the choice set, with absolute one-SD impacts of 4.1 and 4.0 percentage points and completeness values of 21 and 20 percent, respectively. Figure 31 in Appendix E illustrates this pattern: when excess dissimilarity is low, subjects' responses to differences in instrumental value more closely resemble the step function predicted by optimal choice.

Finally, Figure 30 in Appendix E repeats the analysis of Figure 7 under the alternative design, showing that the four representative characteristics selected in the main study retain qualitatively similar impacts.

Result 5 Aversion to larger structures, apparent neglect of signal likelihoods, and attenuated responses to value differences when comparisons are difficult persist across subject subsamples, subject pools, and alternative experimental designs.

5 Discussion

According to standard economic theory, the demand for information is driven by its instrumental value—the extent to which it improves decision-making in relevant tasks. We test this hypothesis and study whether *other* characteristics of information structures also systematically influence demand. To do this, we use a novel experiment that measures subjects' relative preference rankings for hundreds of distinct information structures that differ not only in their instrumental value but also in a number of other formal characteristics. These non-instrumental characteristics plausibly influence the type of information processing required to compare information structures and therefore may shape the difficulty of identifying optimal sources of information.

As predicted by standard theory, we find that instrumental value plays a major role in shaping demand for information structures. However, in 27 percent of cases, subjects prefer less instrumentally valuable structures over more valuable ones, leading to substantial earnings losses. These errors in *assessing* information structures contrast sharply with subjects' near-perfect ability to *use* signals from the same information structures. Our results therefore suggest that even when

impact under the alternative design, and the difference in the probability of maximally certain signals, ΔP(max certain), has a relatively strong impact in the main study. Both patterns could reflect attempts to simplify evaluation by focusing on the likelihood of particular types of signals rather than assessing the information structure as a whole.

information is easy to use (ex post), assessing the relative value of information structures (ex ante) can be difficult and prone to error.

Importantly, we find that characteristics of information structures beyond instrumental value predict preference rankings in systematic ways: subjects are averse to larger structures (such as those that generate more distinct posteriors or more signals with residual uncertainty), they tend to rely on a simplified evaluation that neglects signal likelihoods, and their response to differences in instrumental value is attenuated when the two structures have a low total value or involve tradeoffs that are harder to assess. These patterns persist across subject subsamples, in a laboratory experiment and its online replication, and in an additional experiment that uses a more standard presentation of information structures and a simpler elicitation of information demand. We also show that a small set of representative structure characteristics reflecting these patterns captures most of the explainable variation in preference rankings, with impacts comparable in magnitude to those of the strongest participant-level predictors, such as cognitive scores and confidence.

Taken together, our findings suggest that information demand depends not only on instrumental value but also on how difficult it is to assess and compare value across sources. Evaluating an information structure requires (i) considering all possible signal realizations, (ii) identifying the optimal action conditional on each signal, (iii) computing the associated expected payoffs, and (iv) weighting these payoffs by the likelihoods of the signals to arrive at an overall value. The three patterns we document suggest distinct responses to these cognitive demands: favoring structures that are simpler to evaluate, approximating value through a simpler calculation, and responding less to value differences when comparisons are difficult. Because these responses exhibit systematic structure, our findings offer useful insights into how cognitive constraints shape information demand, suggest a path toward more predictive models of how individuals select and value information structures, and have implications for efficient information design and provision.

Furthermore, we document a counterintuitive relationship between subjects' cognitive sophistication and the extent to which structure characteristics predict their rankings. The representative structure characteristics are *more* predictive of rankings for more sophisticated subjects (e.g., subjects with high cognitive scores or high confidence in their rankings), despite the fact that more sophisticated subjects also make fewer errors in ranking structures.[48] Less sophisticated subjects

[48] In a strategic setting (i.e., repeated p-beauty contest), Gill & Prowse (2022) find subjects are heterogeneous in their responses to the complexity of strategic situations, with higher level-k (more sophisticated) subjects being more likely to respond strongly to complexity. This is broadly parallel to our finding that more sophisticated subjects' demand for information is more systematically shaped by non-instrumental (complexity) characteristics of information structures.

make errors at a higher rate, but they also make errors that are less predictable and plausibly more random, while more sophisticated subjects make errors that are better organized by systematic characteristics of information structures. This may offer insight into how individuals respond to complexity. Recognizing and responding to complexity itself requires effort and insight.[49] Additionally, structure characteristics are more predictive of preference rankings among subjects who are ambiguity averse, providing further evidence of a link between responses to complexity and ambiguity, as highlighted in the recent literature.

Our experimental design deliberately abstracts from other forces that may influence information demand.

First, we present information structures to subjects in a way (using either the partition representation or the more standard tabular format) that makes their formal characteristics—such as the number and likelihood of signals—transparent and their instrumental value relatively intuitive to understand. This is confirmed in our results: the majority of decisions align with instrumental value, allowing us to study systematic deviations in an environment where decision making is otherwise highly optimal.[50] As noted earlier, the way we present information structures also effectively removes errors in the *use* of information. We do not claim that such errors are unimportant for information demand. Rather, our point is that even when subjects can correctly interpret and use signals, identifying the relative value of information structures remains fundamentally challenging. While our design abstracts from many aspects of how information structures are encountered in applied settings, it provides a useful benchmark for isolating the difficulty of evaluating information in a simple setting, much as experiments on risk-taking do by eliciting choices over abstract lotteries.

Second, the deliberately abstract framing of our experiment has both advantages and limitations. On the positive side, this abstraction allows us to tightly control several formal characteristics of information structures, most importantly instrumental value, providing the kind of experimental precision that enables relatively sharp conclusions about the drivers of information demand. At the same time, by design, our experiment omits other forces that may shape information demand in

[49] This perspective resonates with the recent work by Agranov, Schotter & Trevino (2025), who propose that complexity is subjective and demonstrate systematic heterogeneity in how people perceive and engage with decision problems. Our finding that sophisticated subjects make more predictable errors aligns with their framework, as systematic responses to complexity, even erroneous ones, require some degree of cognitive engagement with the problem.

[50] We also focus on the simplest possible decision problem (a guessing task) to study information demand. An interesting direction for future research is to study how the complexity of the decision problem interacts with difficulties in identifying instrumental value.

applied settings, many of which are likely to be context-dependent. For instance, our results cannot speak to how voters might be drawn to information sources that reinforce their ideological identity, or how employees might seek out feedback that preserves self-esteem rather than maximizes learning. These motivations, while important in many real-world environments, operate alongside the cognitive and instrumental considerations we isolate. Extending our methods to richer settings in which such "non-cognitive" factors also play a role, and comparing their influence to the regularities we identify, seems a promising next step in this line of research.

Despite this, the non-instrumental drivers of informational preferences we identify are intuitive and appear relevant to real-world information environments. In many decisions—a physician choosing among diagnostic tests, an investor selecting between forecasting models, a teacher deciding which type of student assessment data to rely on—the decision maker must choose between information sources that differ not only in expected usefulness but also in how complex they are to evaluate. Our findings suggest that decision makers might be drawn to *simple* information sources and characteristics like posterior dispersion, residual uncertainty, and signal asymmetries may shape which information sources are considered simple and more useful. For example, a doctor may choose a diagnostic test with binary results over a probabilistic one not because the latter is harder to act on, but because its value is harder to evaluate in advance. Similarly, a recent literature on AI-assisted decision making shows that *coarsening* algorithmic recommendations not only improves decision quality but also increases demand for such information (see Dreyfuss & Hoong (2025)). Likewise, sources that leave residual uncertainty, such as diagnostic tests or models that fail to produce definitive results, may appear less valuable, not because they convey less information, but because their partial signals are harder to interpret ex ante. Finally, people may favor information sources which appear highly useful when asymmetries in signal probabilities are ignored. For example, an investor comparing forecasting models may focus on how often each model "gets it right" across different market conditions, implicitly weighting all conditions equally even though some occur far more frequently than others. Taken together, these examples illustrate how the systematic patterns we observe in a highly controlled experimental setting can shed light on why decision makers in practice may favor simpler or more transparent sources of information, even when more instrumentally valuable alternatives are available.

References

Agranov, M., Schotter, A. & Trevino, I. (2025), Complex for whom? an experimental approach to subjective complexity, working paper.

Ambuehl, S. & Li, S. (2018), 'Belief updating and the demand for information', *Games and Economic Behavior* **109**, 21–39.

Andries, M. & Haddad, V. (2020), 'Information aversion', *Journal of Political Economy* **128**(5), 1901 – 1939.

Arrieta, G. & Nielsen, K. (2025), Procedural decision-making in the face of complexity, working paper.
URL: *https://garrieta.github.io/jmp/JMP.pdf*

Ba, C., Bohren, J. A. & Imas, A. (2023), Over- and underreaction to information, working paper.

Bouchouicha, R., Oprea, R., Vieider, F. M. & Wu, J. (2025), Is Prospect Theory Really a Theory of Choice. Unpublished manuscript.

Caplin, A. & Leahy, J. (2001), 'Psychological expected utility theory and anticipatory feelings', *The Quarterly Journal of Economics* **116**(1), 55–79.

Charness, G., Oprea, R. & Yuksel, S. (2021), 'How do people choose between biased information sources? evidence from a laboratory experiment', *Journal of the European Economic Association* **19**(3), 1656–1691.

Danz, D., Vesterlund, L. & Wilson, A. J. (2022), 'Belief elicitation and behavioral incentive compatibility', *American Economic Review* **112**(9), 2851–83.

de Clippel, G., Moscariello, P., Ortoleva, P. & Rozen, K. (2024), Caution in the face of complexity, working paper.

Dillenberger, D. (2010), 'Preferences for one-shot resolution of uncertainty and allais-type behavior', *Econometrica* **78**(6), 1973–2004.

Dreyfuss, B. & Hoong, R. (2025), Calibrated coarsening: Designing information for ai-assisted decisions, working paper.

Eliaz, K. & Schotter, A. (2007), 'Experimental testing of intrinsic preferences for noninstrumental information', *American Economic Review* **97**(2), 166–169.

Eliaz, K. & Schotter, A. (2010), 'Paying for confidence: An experimental study of the demand for non-instrumental information', *Games and Economic Behavior* **70**(2), 304–324.

Ely, J., Frankel, A. & Kamenica, E. (2015), 'Suspense and surprise', *Journal of Political Economy* **123**(1), 215–260.

Enke, B. (2024), The cognitive turn in behavioral economics, working paper.

Enke, B. & Graeber, T. (2023), 'Cognitive uncertainty', *Quarterly Journal of Economics* **138**(4), 2021–2067.

Enke, B., Graeber, T. & Oprea, R. (2025), 'Complexity and time', *Journal of the European Economic Association* p. jvaf009.

Enke, B., Graeber, T., Oprea, R. & Yang, J. (2026), 'Behavioral attenuation', *Journal of Political Economy* . Forthcoming.

Enke, B. & Shubatt, C. (2026), Quantifying lottery choice complexity, working paper.

Falk, A. & Zimmermann, F. (2024), 'Attention and dread: Experimental evidence on preferences for information', *Management Science* **70**(10), 7090–7100.

Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Stanford University Press.

Fudenberg, D., Kleinberg, J., Liang, A. & Mullainathan, S. (2022), 'Measuring the completeness of economic models', *Journal of Political Economy* **130**(4), 956–990.

Gill, D. & Prowse, V. (2022), 'Strategic complexity and the value of thinking', *The Economic Journal* **133**(650), 761–786.

Golman, R. & Loewenstein, G. (2018), 'Information gaps: A theory of preferences regarding the presence and absence of information', *Decision* **5**(3), 143–164.

Golman, R., Loewenstein, G., Molnar, A. & Saccardo, S. (2022), 'The demand for, and avoidance of, information', *Management Science* **68**(9), 6454–6476.

Gonçalves, D. (2024), Speed, accuracy and complexity, working paper.

Grant, S., Kajii, A. & Polak, B. (1998), 'Intrinsic preference for information', *Journal of Economic Theory* **83**(2), 233–259.

Greiner, B. (2015), 'Subject pool recruitment procedures: organizing experiments with orsee', *Journal of the Economic Science Association* **1**(1), 114–125.

Grether, D. M. & Plott, C. R. (1979), 'Economic theory of choice and the preference reversal phenomenon', *The American Economic Review* **69**(4), 623–638.

Hossain, T. & Okui, R. (2013), 'The binarized scoring rule', *The Review of Economic Studies* **80**(3 (284)), 984–1001.

Jones, M. & Sugden, R. (2001), 'Positive confirmation bias in the acquisition of information', *Theory and Decision* **50**, 59–99.

Koszegi, B. & Rabin, M. (2009), 'Reference-dependent consumption plans', *American Economic Review* **99**(3), 909–36.

Kreps, D. M. & Porteus, E. L. (1978), 'Temporal resolution of uncertainty and dynamic choice theory', *Econometrica* **46**(1), 185–200.

Liang, Y. (2023), Boundedly rational information demand, working paper.

Lichtenstein, S. & Slovic, P. (1971), 'Reversals of preference between bids and choices in gambling decisions.', *Journal of experimental psychology* **89**(1), 46.

Llorente-Saguer, A., Oliveros, S. & Zultan, R. (2025), Beyond value: on the role of symmetry in demand for information, working paper.

Martinez-Marquina, A., Niederle, M. & Vespa, E. (2019), 'Failures in contingent reasoning: the role of uncertainty', *American Economic Review* **109**(10), 3437–3474.

Masatlioglu, Y., Orhun, Y. & Raymond, C. (2023), 'Intrinsic information preferences and skewness', *American Economic Review* **113**(10), 2615–44.

Montanari, G. & Nunnari, S. (2022), Audi alteram partem: an experiment on selective exposure to information, working paper.

Niederle, M. & Vespa, E. (2023), 'Cognitive limitations: Failures of contingent thinking', *Annual Review of Economics* **15**(Volume 15, 2023), 307–328.

Nielsen, K. (2020), 'Preferences for the resolution of uncertainty and the timing of information', *Journal of Economic Theory* **189**.

Novák, V., Matveenko, A. & Ravaioli, S. (2024), 'The status quo and belief polarization of inattentive agents: Theory and experiment', *American Economic Journal: Microeconomics* **16**(4), 1–39.

Oprea, R. (2024), 'Decisions under risk are decisions under complexity', *American Economic Review* **114**(12), 3789–3811.

Oprea, R. (2026), Complexity and its measurement, *in* A. Rees-Jones, ed., 'Handbook of Experimental Methods in the Social Sciences', Edward Elgar, chapter 3, pp. 104–161.

Oprea, R. & Yuksel, S. (2022), 'Social exchange of motivated beliefs', *Journal of the European Economic Association* **20**(2), 667–699.

Palacios-Huerta, I. (1999), 'The aversion to the sequential resolution of uncertainty', *Journal of Risk and Uncertainty* **18**(3), 249–269.

Puri, I. (2025), 'Simplicity and risk', *The Journal of Finance* **80**(2), 1029–1080.

Shubatt, C. & Yang, J. (2024), Tradeoffs and comparison complexity, working paper.

Zimmermann, F. (2015), 'Clumped or piecewise? evidence on preferences for information', *Management Science* **61**(4), 740–753.

Online Appendix for

Beyond Instrumental Value: How Complexity Shapes Information Demand

Menglong Guan Ryan Oprea Sevgi Yuksel

CONTENTS:

A. Characteristics of Information Structures and Participant-Level Characteristics

B. Additional Plots and Tables

C. Lasso and Decision Tree

D. Lab Study and Replication Study

E. Study with Alternative Information Representation and Preference Elicitation

F. Screenshots and Instructions

A Characteristics of Information Structures and Participant-Level Characteristics

Table 3: Characteristics of Information Structures: Summary Statistics

	Structure characteristic	Min	Max	Mean	Median	Variance
Difference	Instrumental value (ΔV)	0.1	0.4	0.1916	0.2	0.0091
	Number of signals (Δn_s)	-3	4	0.4694	0	1.7164
	Number of signals with residual uncertainty (Δn_{ru})	-4	2	-1.1315	-1	0.9904
	Number of distinct posteriors (Δn_{dp})	-2	3	-0.1234	0	1.0029
	Unweighted guessing accuracy ($\Delta \tilde{V}$)	-0.1778	0.4444	0.161	0.1458	0.0117
	Probability of maximally certain signals (ΔP(max certain))	-0.4	1	0.4222	0.4	0.0811
	Probability of maximally uncertain signals (ΔP(max uncertain))	-0.8	0.6	-0.2228	-0.2	0.0828
	Probability of worse-than-prior signals (ΔP(worse than prior))	-0.9	0.7	-0.2661	-0.2	0.092
	Probability of guessing black (ΔP(guess black))	-0.7	0.5	-0.0631	-0.1	0.0338
	Skewness (Δskewness)	-5.3334	3.0749	-0.5235	-0.3765	1.355
	Informativeness (ΔI)	-0.1146	0.971	0.4197	0.39	0.0564
Sum	Instrumental value ($\sum V$)	0.1	0.7	0.3995	0.4	0.029
	Number of signals ($\sum n_s$)	3	10	7.727	8	1.8095
	Number of signals with residual uncertainty ($\sum n_{ru}$)	1	7	2.8749	3	1.6666
	Number of distinct posteriors ($\sum n_{dp}$)	3	8	5.3328	5	0.8231
	Unweighted guessing accuracy ($\sum \tilde{V}$)	0.0111	0.7667	0.4877	0.5	0.0256
	Probability of maximally certain signals ($\sum$P(max certain))	0	1.8	0.9152	0.9	0.1739
	Probability of maximally uncertain signals ($\sum$P(max uncertain))	0	1.4	0.3276	0.2	0.1052
	Probability of worse-than-prior signals ($\sum$P(worse than prior))	0	1.6	0.3835	0.4	0.1243
	Probability of guessing black ($\sum$P(guess black))	0.7	1.9	1.2917	1.3	0.0355
	Skewness ($\sum$skewness)	-4.1667	3.5396	-0.3035	-0.4636	1.3724
	Informativeness ($\sum I$)	0.0913	1.742	0.9515	0.9768	0.1506
Blackwell order (Blackwell)		0	1	0.8635	1	0.1178
Refine		0	1	0.106	0	0.0948
Excess dissimilarity		0	0.14	0.0075	0	0.0004
Value-dissimilarity ratio		0.4167	1	0.9549	1	0.0133

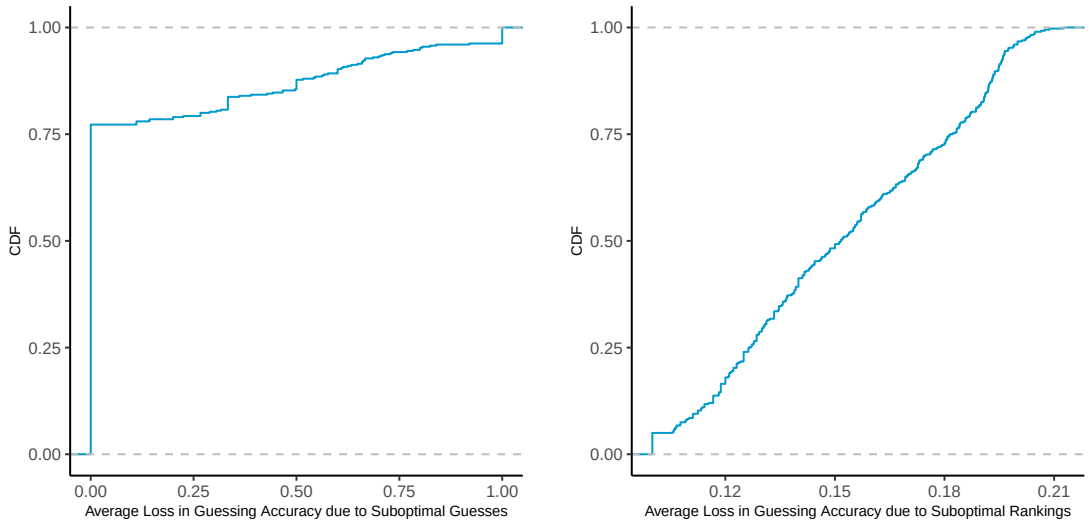
Notes: We focus on the pairwise comparisons of information structures that have a strict difference in instrumental value in our data.

Table 4: Participant Characteristics

Subject characteristic	**Description**
Cognitive score	Overall score on the ten cognitive questions: one Wason selection task, three cognitive reflection test questions, three Raven's test questions, and three belief bias syllogism questions.
Risk	Certainty equivalent to a simple lottery that pays \$30 with a 50% chance.
Ambiguity	Certainty equivalent to an Ellsberg lottery that has the same expected value as the simple lottery.
Compound	Certainty equivalent to a compound lottery that has the same expected value as the simple lottery.
Confidence in ranking optimality	Confidence in the optimality of rankings, measuring how certain subjects are in the optimality of their rankings of information structures.
Computation of information instrumental value	Computation of the chance that a computer, which always makes optimal guesses, will correctly guess the drawn ball color when receiving information from a given information structure. This is essentially the computation of the instrumental value of the given information structure.
Confidence in computation of information instrumental value	Confidence in correctly computing the instrumental value of the given information structure.
Complex Information computation	Whether a subject faces the relatively more complex information structure in the instrumental value computation question.
Computation of compound lottery value	Computation of the chance of winning the prize of the compound lottery.
Confidence in computation of compound lottery	Confidence in correctly computing the winning chance of the compound lottery.
Mistake in computation of information instrumental value	Mistake in the computation of the instrumental value of the information structure.
Mistake in computation of compound lottery value	Mistake in the computation of the winning chance of the compound lottery.
Absolute mistake in computation of information instrumental value	Absolute mistake in the computation of the instrumental value of the information structure.
Absolute mistake in computation of compound lottery value	Absolute mistake in the computation of the winning chance of the compound lottery.
Ambiguity aversion	How much lower the certainty equivalent to the Ellsberg lottery is compared to the certainty equivalent to the simple lottery. This provides a measure of ambiguity attitude.
Aversion to compound lottery	How much lower the certainty equivalent to the compound lottery is compared to the certainty equivalent to the simple lottery. This provides a measure of attitude to compound risk.

Notes: The table presents the full list of the 16 participant characteristics we consider in Section 4.4.

B Additional Plots and Tables



(a) Losses due to Suboptimal Guesses

(b) Losses due to Suboptimal Ranking

Figure 11: Distribution of Losses due to Suboptimal Guesses and Ranking *Notes: In panel (a) losses in likelihood of winning the bonus due to suboptimal guesses in guessing tasks are computed on the individual level. In panel (b) losses in likelihood of winning the bonus due to suboptimal rankings in ranking tasks are computed on the individual level, assuming optimal guesses.*

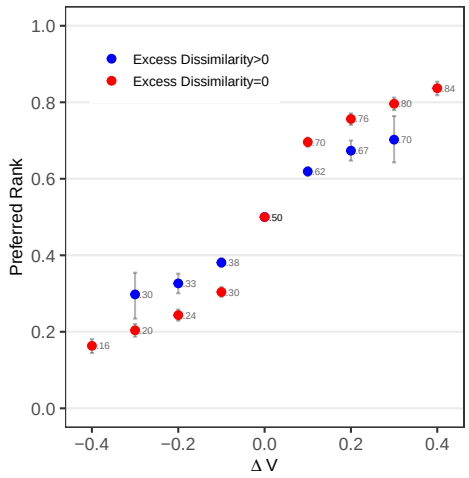


Figure 12: Pairwise Ranking of Information Structures by Value Difference: The Impact of Excess Dissimilarity *Notes: The figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where the two structures present a tradeoff, that is, where excess dissimilarity (defined in Table 1, inspired by Enke & Shubatt (2026)) is strictly positive, with cases where it is zero. In our data, excess dissimilarity is zero in 84 percent of the pairwise comparisons with a strict value difference. The value-dissimilarity ratio (inspired by Shubatt & Yang (2024)) is below one exactly when excess dissimilarity is positive, so the same figure for the value-dissimilarity ratio, with pairwise comparisons split at a ratio of one, is identical and is not shown. Gray lines denote 95 percent confidence intervals by bootstrapping.*

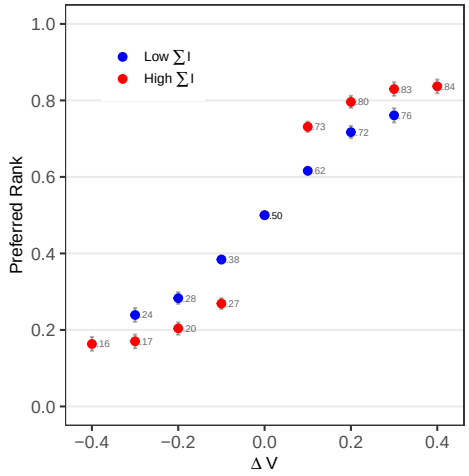


Figure 13: Pairwise Ranking of Information Structures by Value Difference: The Impact of Total Entropy Informativeness *Notes: The figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where the total entropy informativeness of the two structures is either strictly above the median or weakly below. Entropy informativeness of an information structure is defined as the expected reduction of entropy (i.e., a measure of uncertainty in beliefs) from prior to posterior beliefs that are induced by the structure. Gray lines denote 95 percent confidence intervals by bootstrapping.*

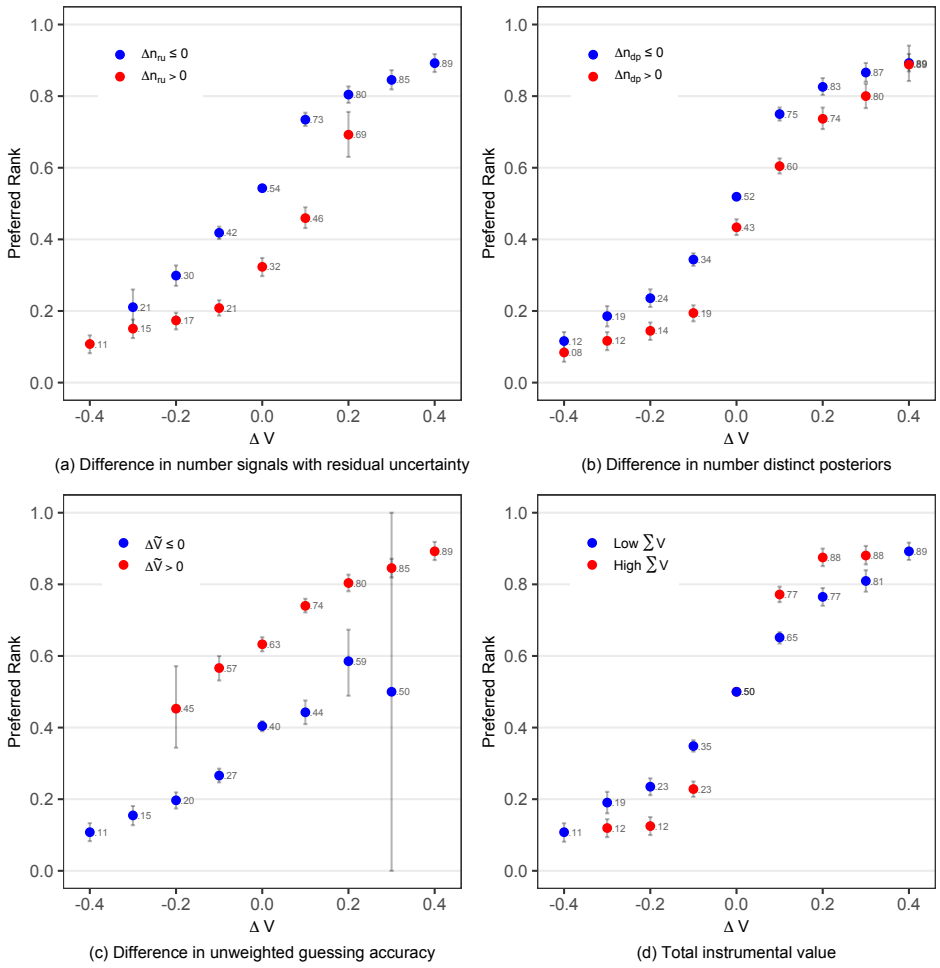


Figure 14: Pairwise Ranking of Information Structures by Value Difference: All Comprehension Questions Answered Correctly at First Try *Notes: This figure replicates Figure 7 but focuses on only subjects who answered all comprehension questions correctly at their first try (153 subjects). Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

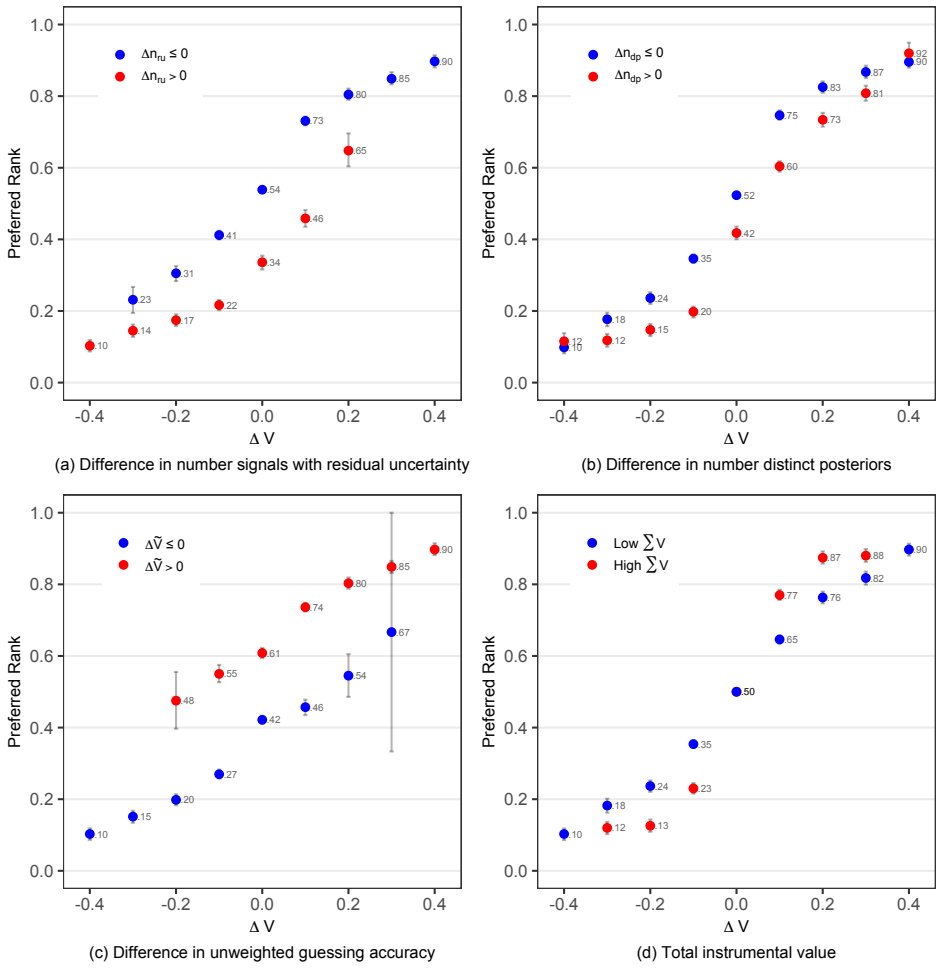


Figure 15: Pairwise Ranking of Information Structures by Value Difference: Guesses Are Always Optimal

Notes: This figure replicates Figure 7 but focuses on only subjects who make optimal guesses all the time (309 subjects). Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.

Table 5: Impacts of Structure Characteristics on Ranking Preference

	Ranking Preference (Logit)				
	(1)	(2)	(3)	(4)	(5)
Difference in number of signals with residual uncertainty (Δn_{ru})	-0.420***				-0.082**
	(0.027)				(0.034)
Difference in number of distinct posteriors (Δn_{dp})		-0.285***			-0.144***
		(0.019)			(0.021)
Difference in unweighted guessing accuracy ($\Delta \tilde{V}$)			2.973***		2.917***
			(0.216)		(0.319)
Total instrumental value ($\sum V$)				1.679***	1.297***
				(0.114)	(0.123)
Constant	0.571***	0.990***	0.552***	0.354***	-0.040
	(0.034)	(0.045)	(0.030)	(0.042)	(0.049)
No. of subjects	400	400	400	400	400

*Notes: Analysis is on pairwise comparisons where one structure has a strictly higher instrumental value than the other ($\Delta V > 0$). The dependent variable is whether the structure with a strictly higher instrumental value is ranked as more preferred than the other. Standard errors (clustered at the subject level) in parentheses. *** 1%, ** 5%, * 10% significance.*

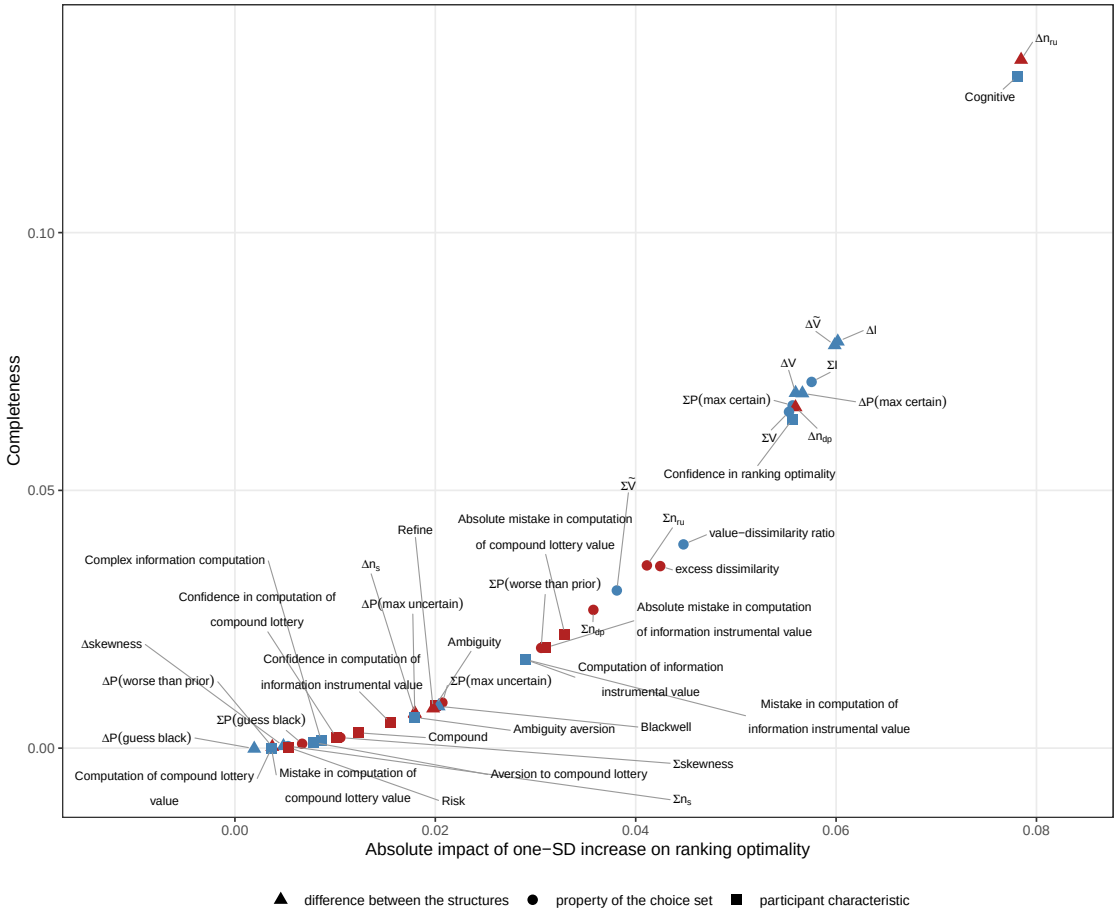


Figure 16: Impact and Completeness of Each Structure and Participant Characteristic Against a Common Benchmark *Notes: Each point is one of the 26 structure characteristics or one of the 16 participant characteristics. The horizontal axis is the absolute value of the impact of a one-standard-deviation increase in the characteristic alone on ranking optimality, as in Figures 6 and 8. The vertical axis is the completeness of a model that adds the characteristic alone to the baseline, where the irreducible error is given by a neural network trained on the raw partitions of the two structures, the 26 structure characteristics, and the 16 participant characteristics together, with the same folds and settings as the benchmarks of Sections 4.3 and 4.4. Triangles denote differences between the two structures being compared or the two comparison relationships, circles denote properties of the choice set, and squares denote participant characteristics. Blue color denotes a positive one-SD impact and red a negative impact. Both analyses focus on all pairwise comparisons with a strict value difference.*

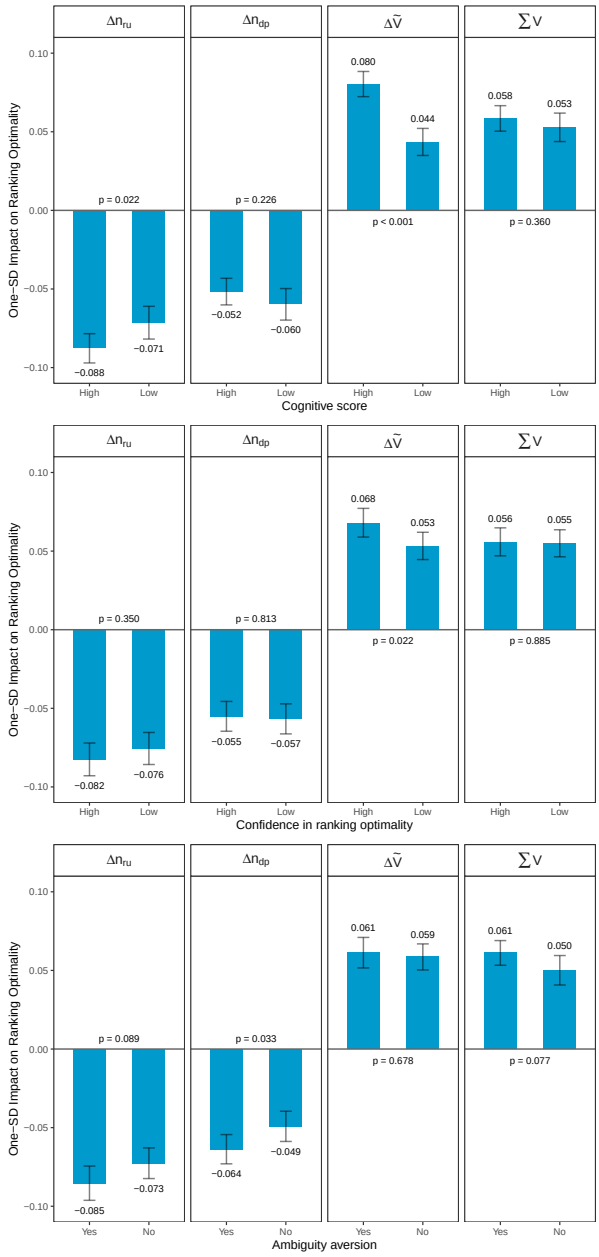


Figure 17: Impact of Structure Characteristics by Subject Type: All Pairwise Comparisons *Notes: Analysis is on all pairwise rankings where one structure is of strictly higher value than the other. Each bar is the impact of a one-standard-deviation increase in the structure characteristic on ranking optimality, estimated separately for subjects whose cognitive score (top panel) or confidence in ranking optimality (middle panel) is strictly above the median value in our sample (High) and for those for whom it is weakly below (Low), and, in the bottom panel, for subjects who are ambiguity averse (Yes) and for those who are not (No). Gray lines denote 95 percent confidence intervals, with standard errors clustered by subject. p-values are for the test that the impacts in the two groups are equal.*

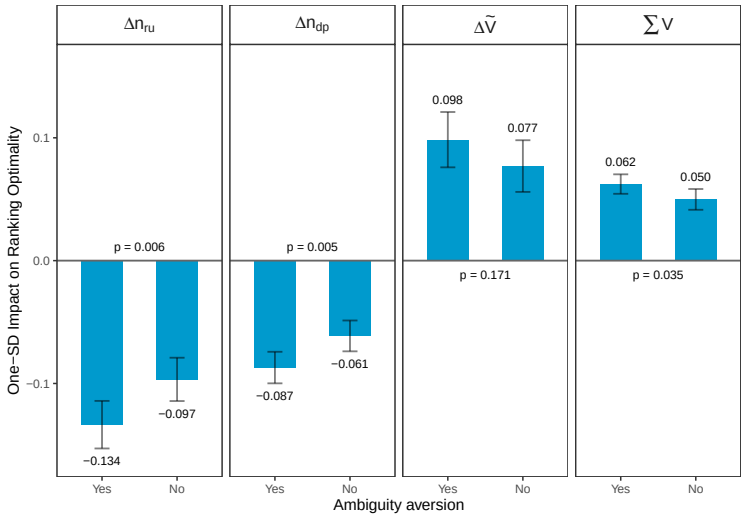


Figure 18: Impact of Structure Characteristics by Ambiguity Attitude *Notes: Analysis is on pairwise rankings where* $\Delta V = 0.1$. *Each bar is the impact of a one-standard-deviation increase in the structure characteristic on ranking optimality, estimated separately for subjects who are ambiguity averse (Yes), that is, whose certainty equivalent for the Ellsberg lottery is lower than that for the simple lottery, and for the other subjects (No). Gray lines denote 95 percent confidence intervals, with standard errors clustered by subject. p-values are for the test that the impacts in the two groups are equal.*

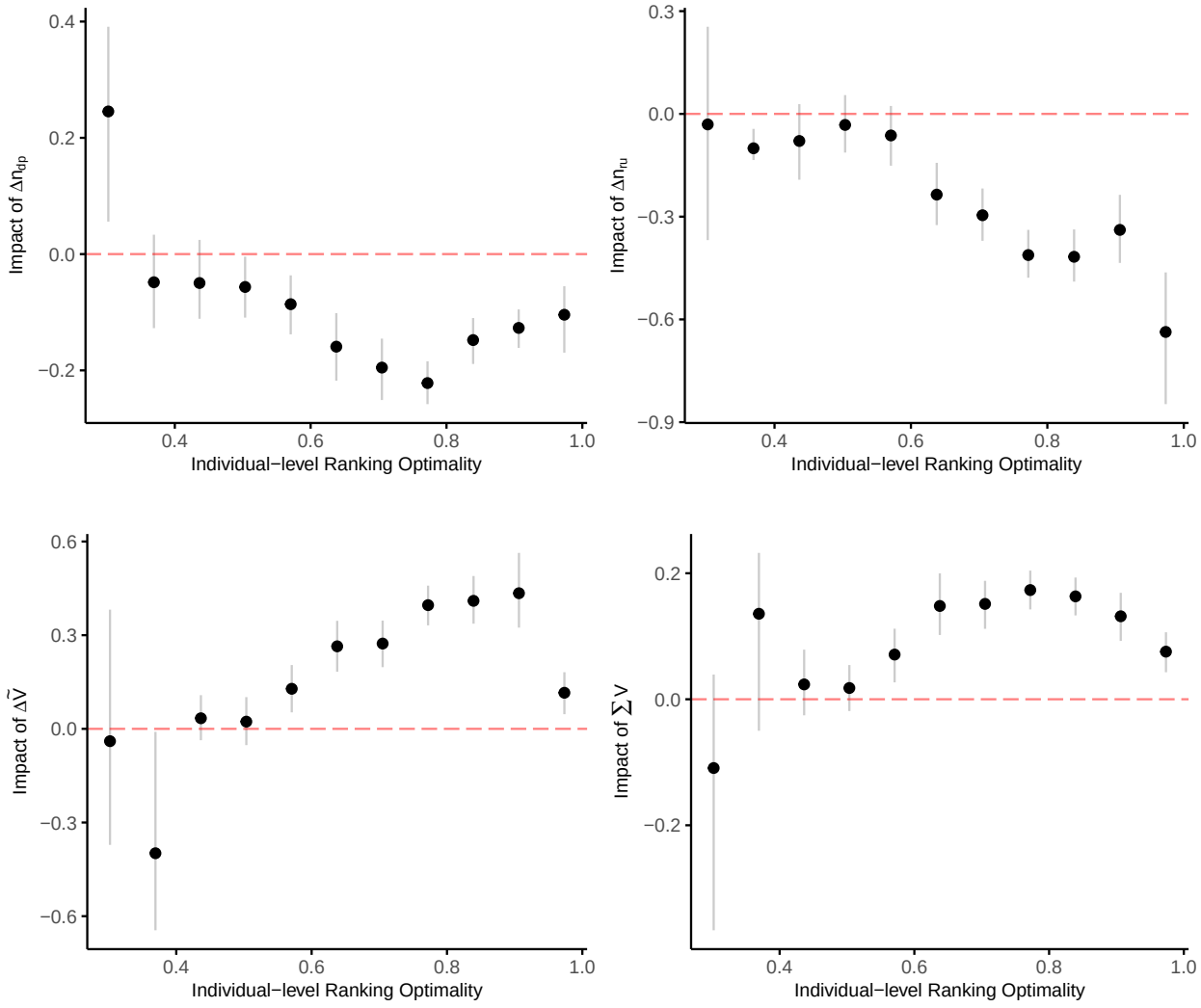


Figure 19: Impact of Selected Characteristics as a Function of Ranking Optimality Rate *Notes: Bin-scatter plots of the impacts of four identified structure characteristics across subjects. Analysis is on pairwise rankings where* $\Delta V = 0.1$*. Values (on the y-axis) represent the difference in ranking optimality when characteristic* f *is strictly above the median value in our sample versus weakly below, separated by subjects' overall optimality rate in this sample (on the x-axis). Gray lines denote 95 percent confidence intervals by bootstrapping.*

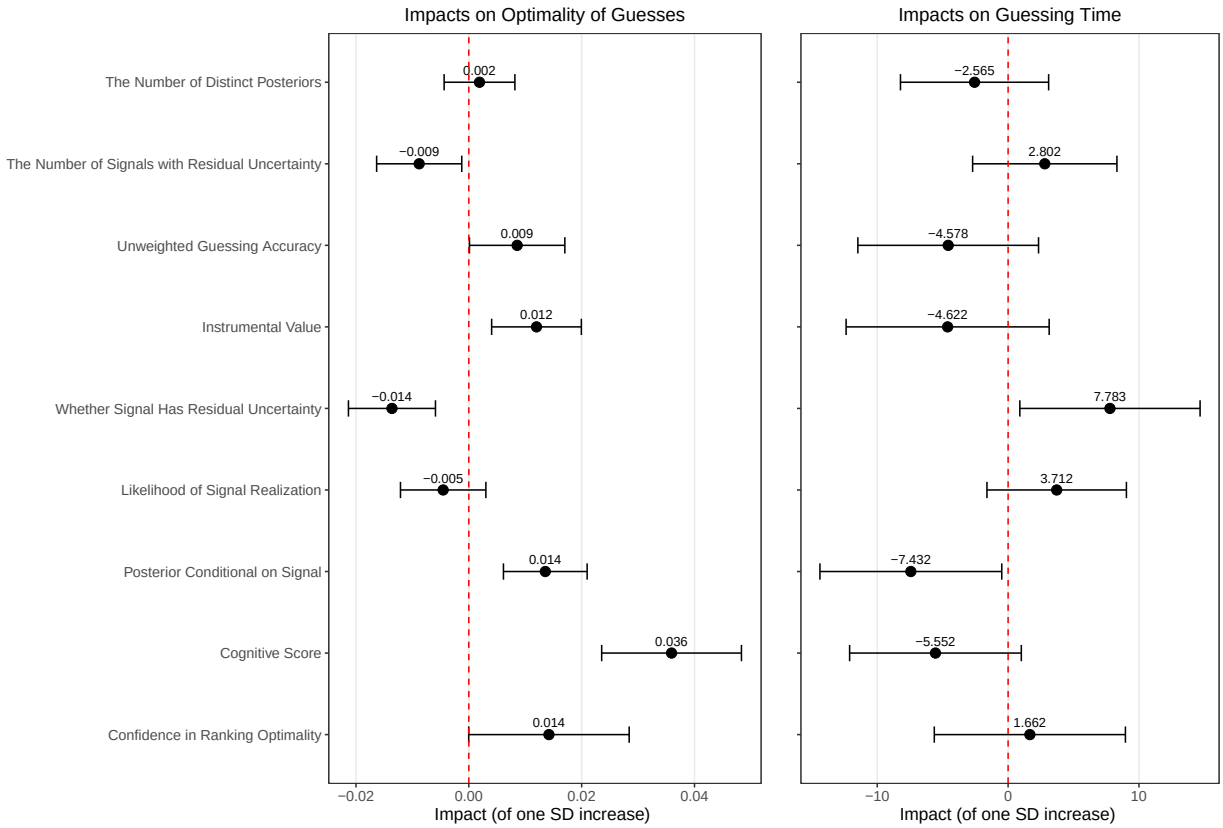


Figure 20: Impact of Structure Characteristics, Signal Characteristics, and Participant Characteristics on Guessing Optimality and Time *Notes: The figure plots the absolute impact of one standard deviation increase in a structure characteristic, signal characteristic, or participant characteristic on guessing optimality and time (in 1/10 seconds). We focus on all pairwise comparisons where one structure is of strictly higher value than the other. Short lines denote 95 percent confidence intervals.*

Table 6: Interaction Effects between Structure Characteristics and Ranking Task Complexity on Ranking Preference

	Ranking Preference (Logit)			
	(1)	(2)	(3)	(4)
Difference in number of distinct posteriors (Δn_{dp})	-0.125***			
	(0.021)			
Δn_{dp} × (Number of fully revealing structures - 1)	-0.080**			
	(0.038)			
Difference in number of signals with residual uncertainty (Δn_{ru})		-0.357***		
		(0.029)		
Δn_{ru} × (Number of fully revealing structures - 1)		-0.071		
		(0.045)		
Difference in unweighted guessing accuracy ($\Delta \tilde{V}$)			3.196***	
			(0.251)	
$\Delta \tilde{V}$ × (Number of fully revealing structures - 1)			0.557*	
			(0.306)	
Total instrumental value ($\sum V$)				0.501***
				(0.128)
$\sum V$ × (Number of fully revealing structures - 1)				0.334**
				(0.138)
Constants	0.788***	0.481***	0.292***	0.578***
	(0.039)	(0.034)	(0.030)	(0.050)
No. of subjects	400	400	400	400

Notes: The seven information structures in each ranking task are randomly selected without replacement while ensuring each task contains information structures of all possible instrumental values $V \in \{0, 0.1, 0.2, 0.3, 0.4\}$ representing the increase in guessing accuracy. Thus, each ranking task contains at least one information structure that fully reveals the state (i.e., $V = 0.4$). Comparisons between a structure that fully reveals the state and another structure should be straightforward, as it is intuitively easy to identify that the former is more valuable. Therefore, ranking tasks containing more structures that fully reveal the state are arguably less complex, given that there are more pairwise comparisons with a structure that fully reveals the state. We use the number of structures that fully reveal the state (out of seven) as a measure of the simplicity *of a ranking task to study whether this measure influences the main results (i.e., the impacts of identified structure characteristics on ranking preference) reported in the paper. Each regression focuses on pairwise comparisons where one structure has a strictly higher instrumental value than the other ($\Delta V > 0$) and the more valuable structure does not fully reveal the state. The dependent variable is whether the structure with a strictly higher instrumental value is ranked as more preferred than the other. Standard errors (clustered at the subject level) in parentheses. *** 1%, ** 5%, * 10% significance.*

Table 7: Impacts of Structure Characteristics on Ranking (Rank-ordered Logit)

	Ranking (1=most preferred, 7=least preferred)					
	(1)	(2)	(3)	(4)	(5)	(6)
Instrumental value (V)	-5.167***	-5.156***	-2.992***	-3.768***	-5.167***	-3.029***
	(0.255)	(0.267)	(0.231)	(0.240)	(0.255)	(0.248)
Number of distinct posteriors (n_{dp})		0.352***				0.302***
		(0.018)				(0.019)
Number of signals with residual uncertainty (n_{ru})			0.401***			0.196***
			(0.029)			(0.034)
Unweighted guessing accuracy ($\tilde{V}$)				-1.703***		-1.306***
				(0.273)		(0.327)
$(\sum_{i=1}^{7} V$ - median$(\sum_{i=1}^{7} V)) \times V$					-0.775	-0.068
					(0.677)	(0.712)
No. of Subjects	400	400	400	400	400	400
No. of Ranking Tasks	3200	3200	3200	3200	3200	3200

*Notes: The unit of analysis is a ranking task. The dependent variable is the ranking of an information structure within the set of seven structures from the same ranking task (1=most preferred, 7=least preferred). A negative (positive) coefficient indicates that structures with higher values of the corresponding characteristic tend to be ranked as more (less) preferred. As discussed in Section 4.3, our primary pairwise comparison analysis examines four identified characteristics: differences in n_{dp}, n_{ru}, $\tilde{V}$, and the sum of V for the two structures being compared. In rank-ordered logit regressions, corresponding to the first three, we directly examine n_{dp}, n_{ru}, and $\tilde{V}$ of each structure. For the fourth characteristic, we use the sum of V of all seven structures in the same ranking task (centered at the median of this measure) and explore whether a lower (higher) $\sum_{i=1}^{7} V$ attenuates (enhances) the impact of instrumental value. Standard errors (clustered at the subject level) in parentheses. *** 1%, ** 5%, * 10% significance.*

C LASSO and Decision Tree

When identifying which information structure characteristics are important predictors of information demand using LASSO and Decision Tree models, we focus on the dataset of pairwise comparisons where, in each pair (denoted as structures A and B), structure A has a strictly higher instrumental value than B.

In LASSO analysis, we fit logistic LASSO regressions with the dependent variable being whether structure A was ranked as more preferred than B, and the independent variables being the 26 structure characteristics. We use a grid of regularization parameter (λ) values, specifically $\log \lambda \in \{-15, -14.9, -14.8, ..., -2.2, -2.1, -2\}$. A larger λ indicates a higher strength of the regularization penalty and will result in a simpler model, i.e., more structure characteristics have zero coefficients. In decision tree analysis, we fit classification decision trees using the Recursive Partitioning and Regression Trees (rpart) algorithm for a grid of tuning parameters $(minsplit, minbucket, cp)$: $\{5, 10, ..., 45, 50\} \times \{3, 5, ..., 17, 19\} \times \{0.01, 0.0075, 0.005, 0.0025, 0.001, 0.00075, 0.0005, 0.00025, 0.0001, 0.00005\}$, where *minsplit* controls the minimum node size for splitting, *minbucket* defines the minimum number of observations that are allowed in a terminal node, and *cp* controls the complexity of a tree. In our analysis, *cp* essentially determines structure characteristic selection in a decision tree.

Table 8a presents the structure characteristics with non-zero coefficients in the LASSO regressions at different values of λ, while Table 8b presents the structure characteristics included in decision trees at different values of *cp*. Notably, Δn_{ru}, $\Delta \tilde{V}$, and Δn_{dp} are selected even in very simple models under both LASSO and decision tree analyses. Although $\sum V$ is selected in slightly more complex models, $\sum I$, which is highly correlated with $\sum V$, is selected early under both analyses.[51] We focus on $\sum V$ rather than $\sum I$ as the former has a more intuitive interpretation.

[51] In decision tree models, Δn_{ru}, $\Delta \tilde{V}$, Δn_{dp} and $\sum I$ appear at the top of the trees (as early splits), meaning they present the strongest predictors.

Table 8: Summary of Structure Characteristic Selection: LASSO and Decision Tree

(a) LASSO

$\log\lambda$	No.	Characteristics with a non-zero coefficient
-2 to -2.5	0	
-2.6 to -3	1	$\boldsymbol{\Delta n_{ru}}$
-3.1 to -3.4	2	$\boldsymbol{\Delta n_{ru}}$, $\sum I$
-3.5 to -3.7	5	$\boldsymbol{\Delta n_{ru}}$, $\sum I$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$
-3.8	6	$\boldsymbol{\Delta n_{ru}}$, $\sum I$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio
-3.9 to -4	7	$\boldsymbol{\Delta n_{ru}}$, $\sum I$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$
-4.1 to -4.6	8	$\boldsymbol{\Delta n_{ru}}$, $\sum I$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$, $\sum n_{dp}$
-4.7	9	$\boldsymbol{\Delta n_{ru}}$, $\sum I$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$, $\sum n_{dp}$, $\sum n_s$
-4.8 to -4.9	10	$\boldsymbol{\Delta n_{ru}}$, $\sum I$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum n_{ru}$, Value-dissimilarity ratio, $\sum \boldsymbol{V}$, $\sum n_{dp}$, $\sum n_s$, Δn_s
...	...	...
-11.1 to -15	26	All 26 characteristics.

(b) Decision Tree

cp	No.	Characteristics, in the order of their first appearance in the tree
{0.01, 0.0075}	0	
{0.005, 0.0025}	3	$\boldsymbol{\Delta n_{ru}}$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$
{0.001, 0.00075}	6	$\boldsymbol{\Delta n_{ru}}$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum I$, $\sum n_s$, ΔP(max certain)
0.0005	10	$\boldsymbol{\Delta n_{ru}}$, $\boldsymbol{\Delta \tilde{V}}$, $\boldsymbol{\Delta n_{dp}}$, $\sum I$, $\sum n_s$, ΔP(max certain), Value-dissimilarity ratio, Δ Skewness, $\sum n_{ru}$, ΔP(max uncertain)
...	...	...
0.00005	26	All 26 characteristics.

Notes: Panel (a) lists, for each band of $\log\lambda$*, the characteristics with a non-zero coefficient, until ten are in. Panel (b) lists, for each value of cp, the characteristics used by the tree, in the order of their first appearance from the top, for the values of cp at which the selection is the same for every combination of minsplit and minbucket. The four representative characteristics of Section 4.3 (*Δn_{ru}*,* Δn_{dp}*,* $\Delta \tilde{V}$*,* $\sum V$*) are in bold in both panels.*

D Laboratory Study and Replication Study

D.1 Details of Laboratory Study

We study a total of 16 information structures, depicted in Figure 21, which were selected to independently vary instrumental value and informativeness (as well as other characteristics). For each of these information structures we (i) study how people make use of these information structures (by eliciting guesses for each information structure and signal) and (ii) study how people value these information structures (by eliciting ranking over the 16 structures and willingness-to-pay for these structures).[52,53]

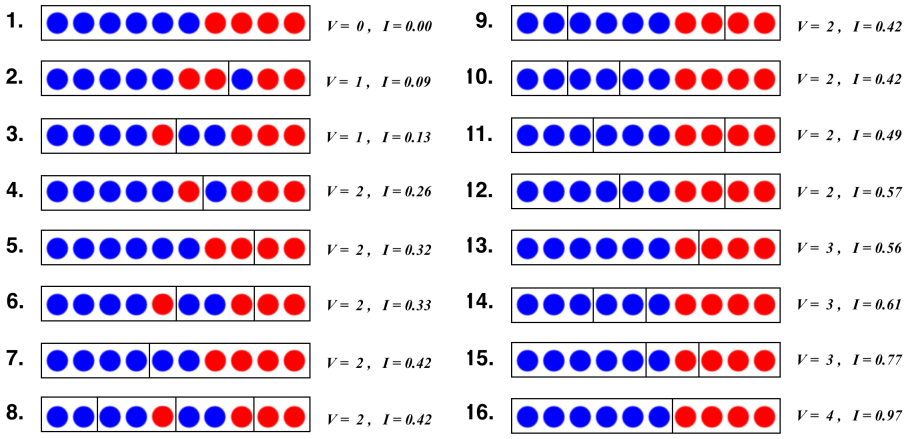


Figure 21: The 16 Information Structures used in the Laboratory Study *Notes: These 16 information structures are a subset of the 237 structures in the Main study. The information structures are listed lexicographically in order of instrumental value (V) and entropy informativeness (I). To make the partitions more salient in the experiment, each cell of the partition was labeled as "Group 1", "Group 2," etc.*

D.2 Treatment Variations in Laboratory Study

In our **Baseline** treatment ($N = 108$ subjects), we ask subjects to perform the Demand section of the experiment (the Ranking and WTP elicitations) first, and the Guesses section afterwards. To this we add two diagnostic treatments that will help us to interpret our results.

First, in our **Reverse** treatment ($N = 53$ subjects), we reversed the order: subjects were assigned Guesses first and Demand afterwards. In these sessions, during the Demand section, subjects were actually shown the guesses they had made earlier, conditional on each possible signal for that information structure.[54] This information was not binding, but was designed to remind

[52] See Figure 32 in Appendix F for a screenshot of Guessing tasks in the laboratory study.
[53] See Figure 33 in Appendix F for screenshots of Ranking and WTP elicitation tasks in the laboratory study.
[54] See Figure 34 in Appendix F for a screenshot of how this was displayed to subjects.

subjects of how different signal realizations are likely to generate different guessing patterns. The purpose of this treatment was to study whether having already made use of information structures (and being reminded of how they are used) improves subjects' identification of instrumental value in guiding their demand. This is useful for understanding the mechanism behind our results.

Second, in our **No Uncertainty** treatment ($N = 61$ subjects), we removed uncertainty from the design using techniques employed by Martinez-Marquina, Niederle & Vespa (2019) and Oprea (2024). Specifically, instead of drawing only one ball, we informed subjects that we would draw *all balls* and pay subjects based on the accuracy of their guess for each of the balls.[55,56] Everything else in the experiment remains identical. Doing this retains much of the information processing involved in evaluating and interpreting information structures, but removes scope for risk, uncertainty or timing preferences (e.g., preferences for the timing of the resolution of uncertainty). Again, this data is useful for understanding the mechanism behind our results.

Implementation Details

All sessions were conducted at the LITE laboratory at the University of California, Santa Barbara between December 2021 and March 2023. We recruited participants from across the curriculum to participate in 15 sessions using the ORSEE recruiting software (Greiner 2015). The experiment was conducted using software programmed by the authors in Qualtrics. Between 8 and 21 participants participated in each session and sessions lasted for 30-40 minutes.

All participants received a show-up fee of $7. Participants' earnings depended on a randomly selected section of the experiment. If the Demand section was selected for payment, we randomized between Ranking and WTP. If Ranking was selected, two information structures (from the set of 16) were randomly selected and the participant was assigned to receive information from the one that was ranked higher. If WTP was selected, one information structure (from the set of 16) was randomly selected and whether or not the participant received information from this source was determined according to the BDM mechanism. In either case (regardless of whether Ranking or WTP was selected), at the end of the experiment, the participant was presented with one additional guessing task with information from the selected information structure and participants were paid based on this guess.[57] If the Guesses section was selected for payment, a random information

[55] In this case, the partition associated with an "information" structure constrains the types of guesses subjects can make by requiring them to make the same guess for all balls in the same cell of the partition.

[56] To achieve a clean comparison to the other treatments, subjects received a prize of $1 for each of the balls for which their guesses were correct. This implies, for example, an information structure that enables a guessing accuracy of 90 percent will generate the same expected bonus payment in all three treatments.

[57] Note that this could mean the participant receives no information depending on their WTP if WTP is selected.

structure was picked and the participant's guesses for that case were used to determine payment for a randomly drawn ball. Note that in all cases, whether or not participants received a bonus payment of $10 ultimately depended on the accuracy of their guess about the color of a randomly selected ball.

D.3 Replication Study

We also ran a **Replication** study on Prolific ($N = 50$ subjects) which used the same 16 information structures from the laboratory study but used the interface of the main study. Our aim was to test whether results would be robust across different subject pools and experimental interfaces. The study was conducted in September 2024. The median time subjects spent was about 17 minutes. Figure 23 replicates Figure 7 with the replication data.

D.4 Results

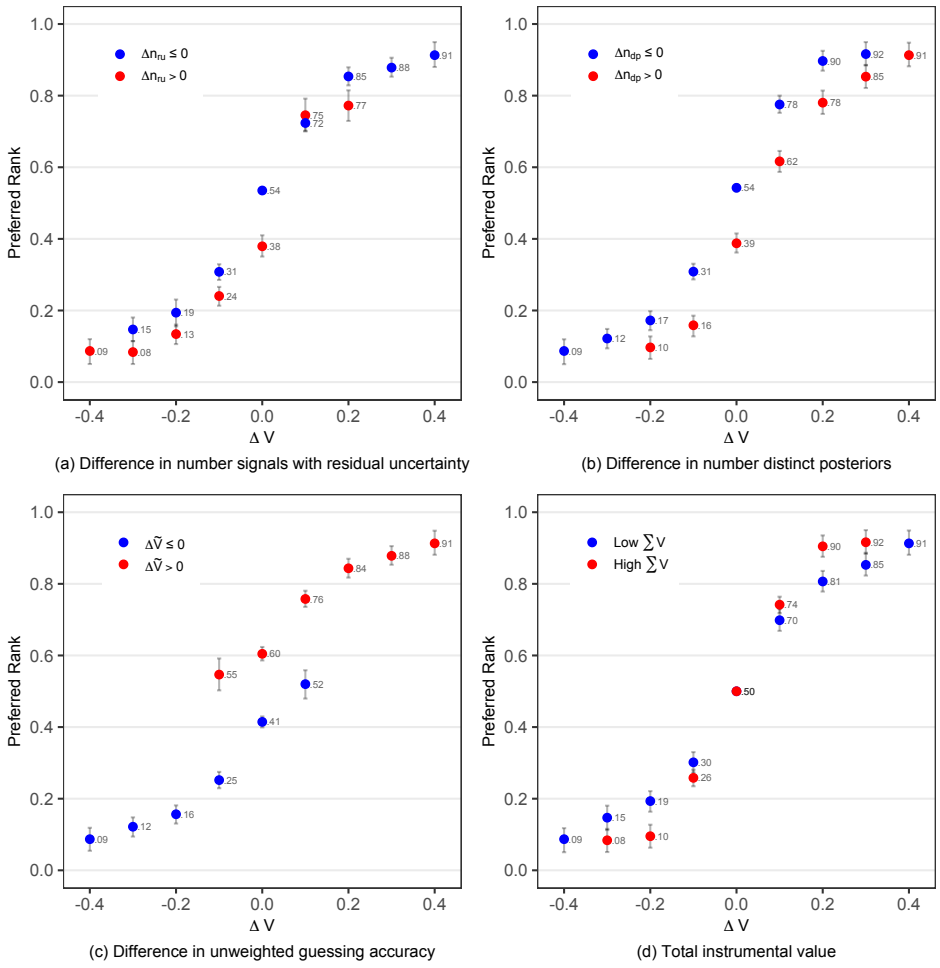


Figure 22: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study *Notes: Both Baseline (Guess Second) and Reverse (Guess First). We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

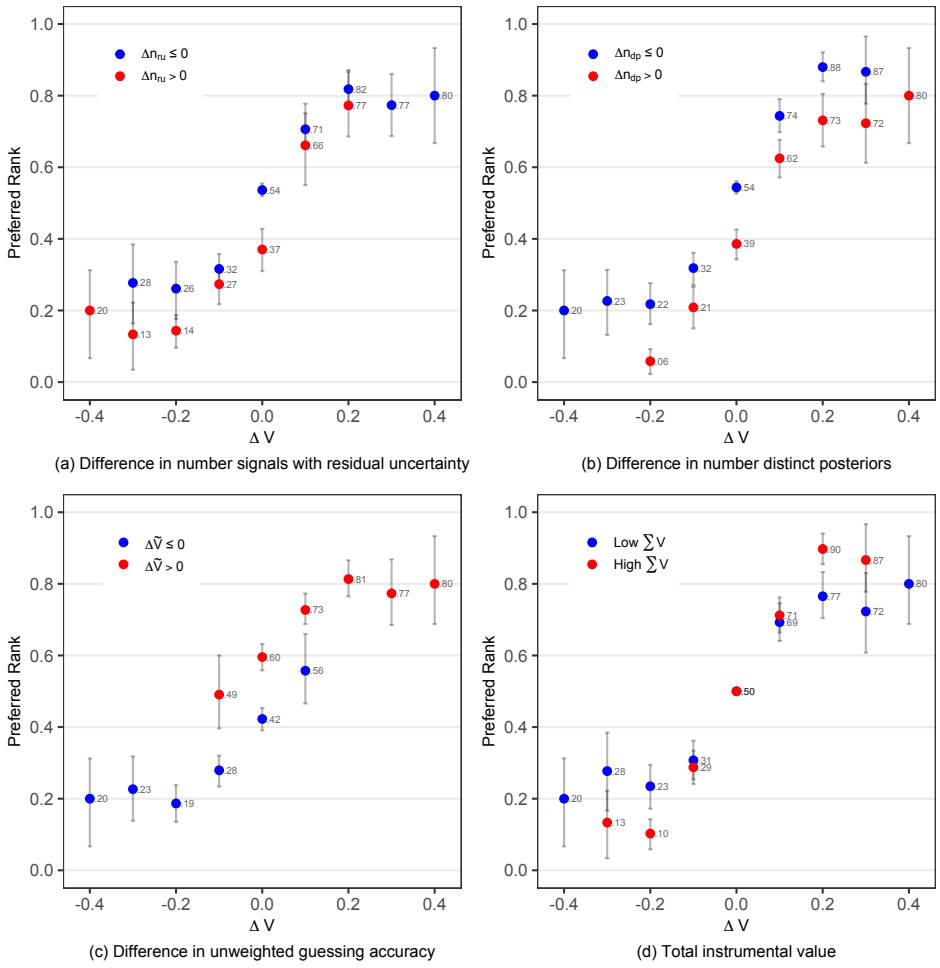


Figure 23: Pairwise Ranking of Information Structures by Value Difference in the Replication Study *Notes: Replication study uses the same experimental interface and subject pool as the main study but focuses on the 16 information structures from the laboratory study. We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

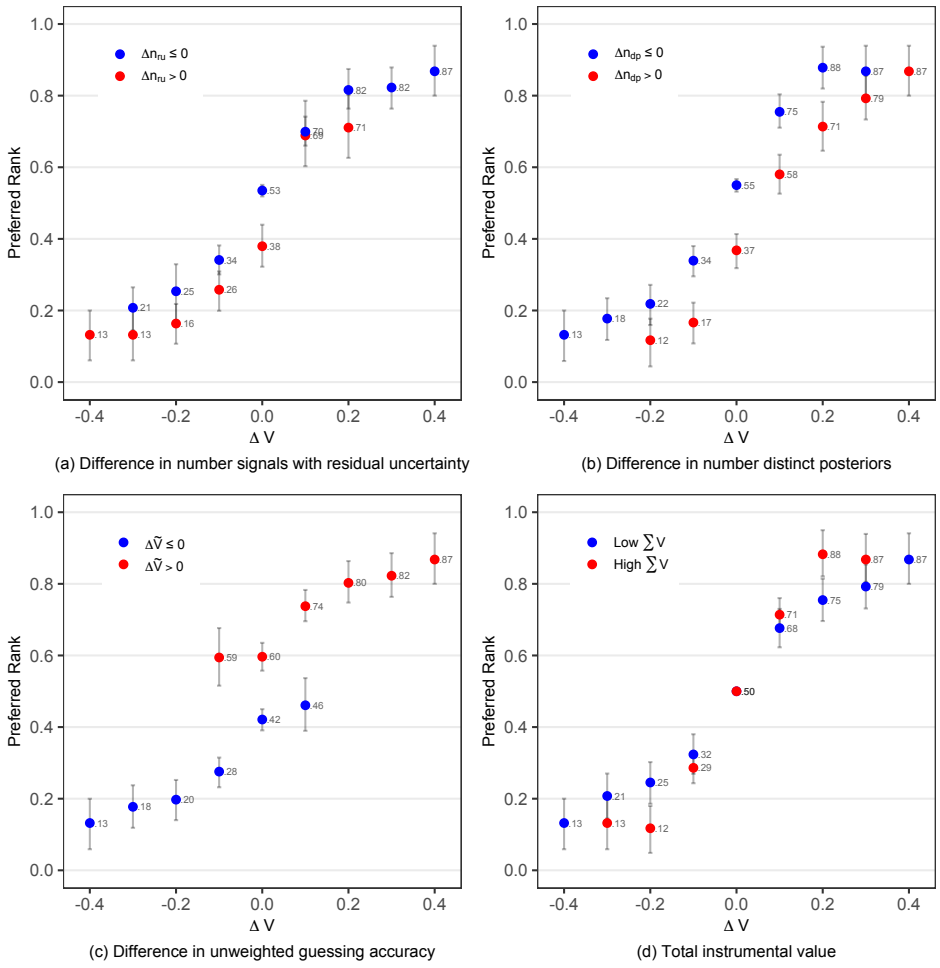


Figure 24: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study (Guess First) *Notes: Subset of subjects in the laboratory study who made guesses first (before elicitation for demand for information). We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

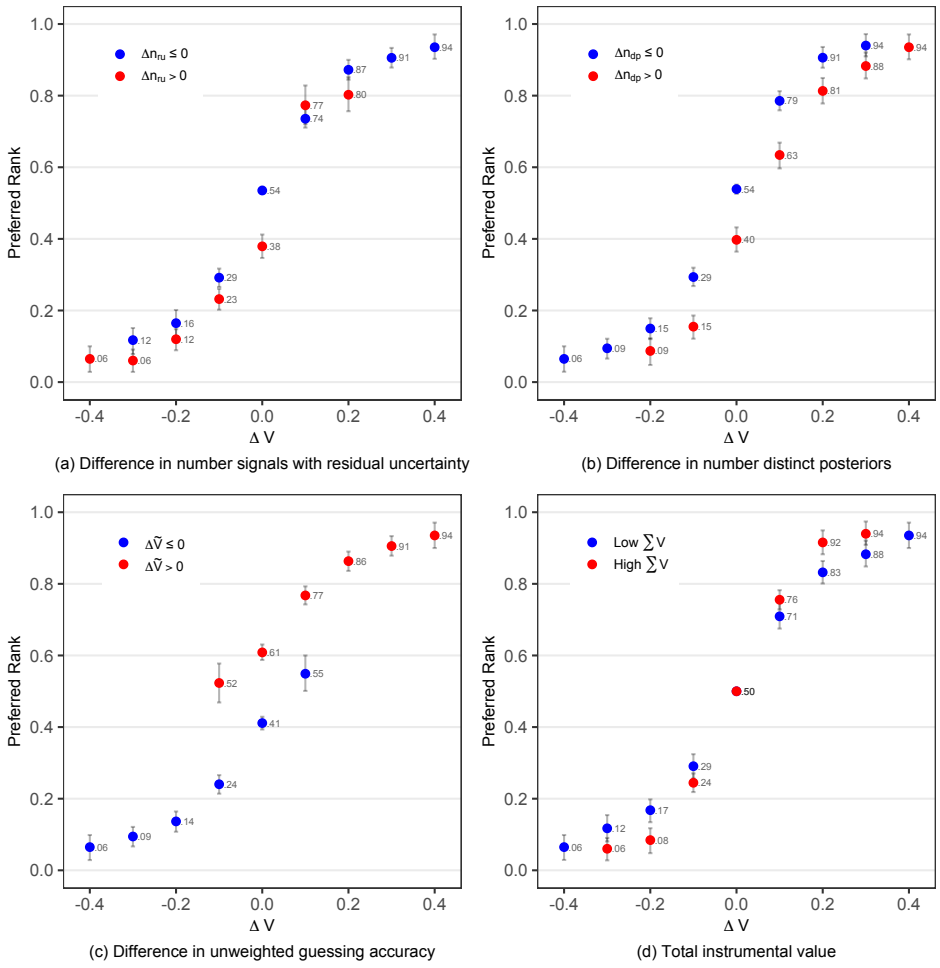


Figure 25: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study (Guess Second) *Notes: Subset of subjects in the laboratory study who made guesses second (after elicitation for demand for information). We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

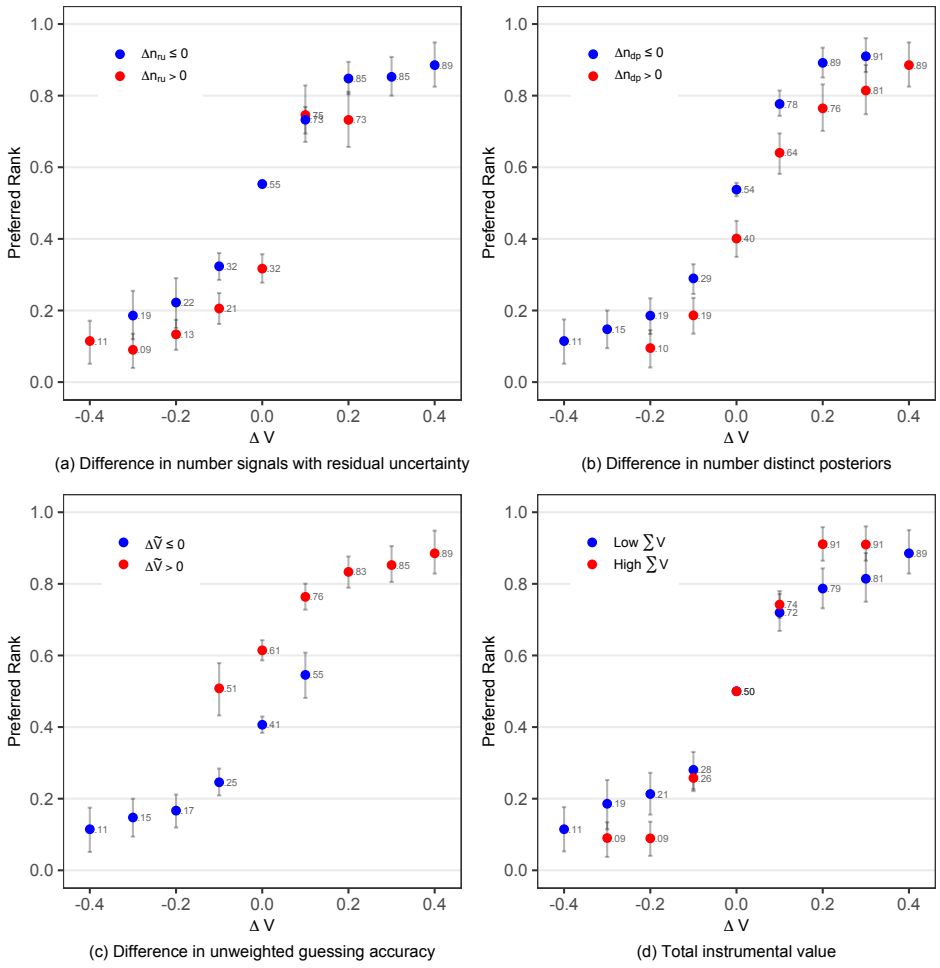


Figure 26: Pairwise Ranking of Information Structures by Value Difference in the Laboratory Study (No Uncertainty Treatment) *Notes: The No Uncertainty treatment of the laboratory study. We focus on pairwise comparisons. Each figure plots the rate at which one structure was preferred to another by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

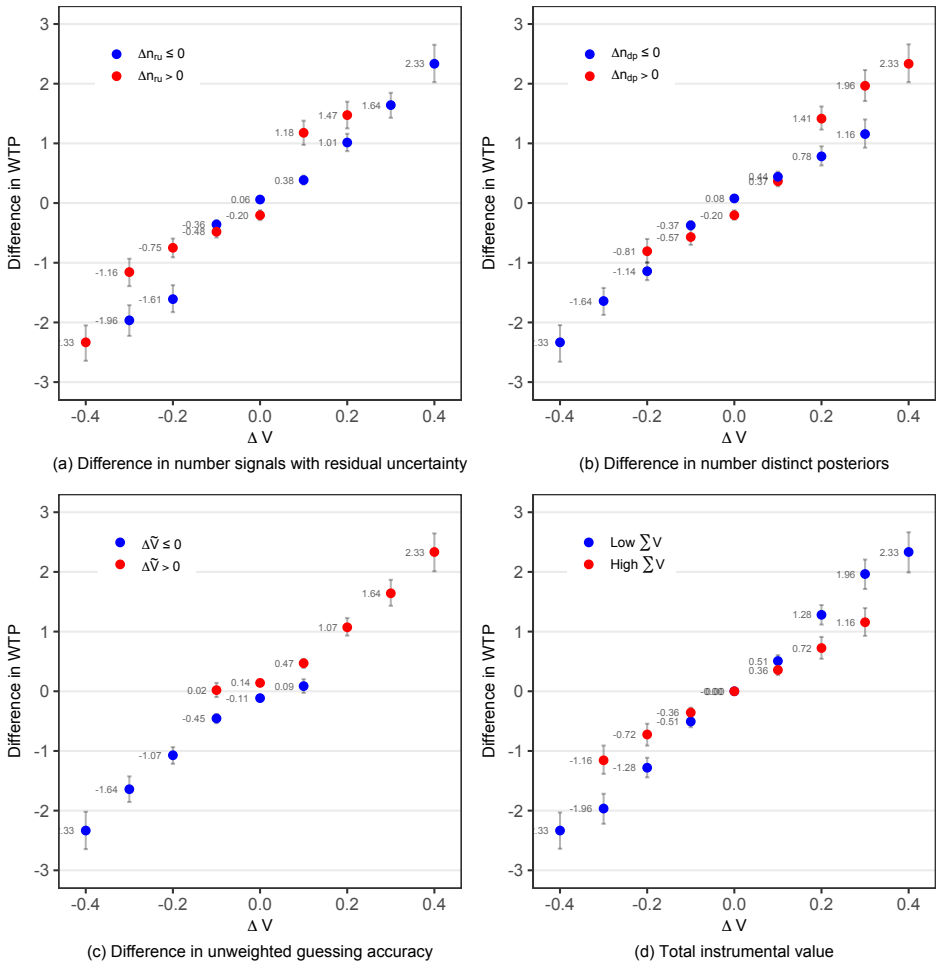


Figure 27: Difference in WTP for Pairwise Information Structures by Value Difference in the Laboratory Study *Notes: Both Baseline (Guess Second) and Reverse (Guess First). We focus on pairwise comparisons. Each figure plots the difference in WTP between a pair of information structures by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

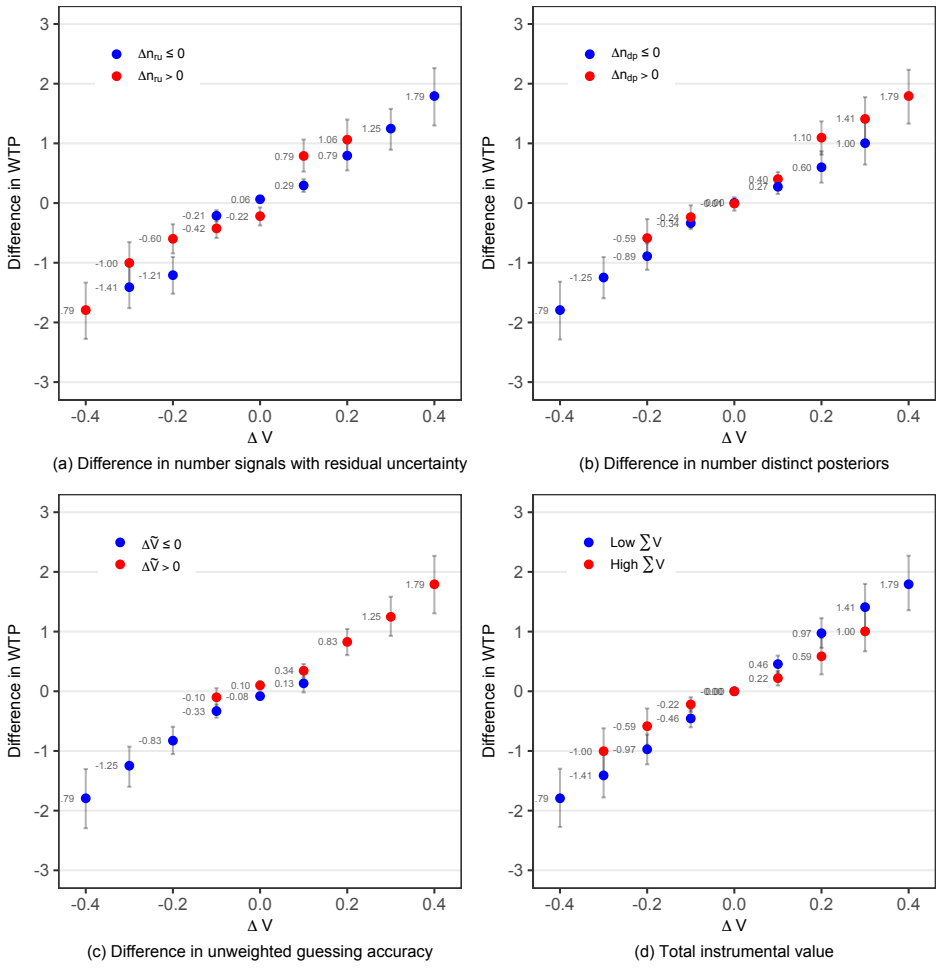


Figure 28: Difference in WTP for Pairwise Information Structures by Value Difference in the Laboratory Study (No Uncertainty Treatment) *Notes: The No Uncertainty treatment of the laboratory study. We focus on pairwise comparisons. Each figure plots the difference in WTP between a pair of information structures by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

E Study with Alternative Information Representation and Preference Elicitation

E.1 Details of the Alternative Design

The study with the alternative design was conducted on Prolific in June and July 2026 with 200 participants. It uses the same universe of 237 information structures but changes the representation of information structures and the format in which preferences over them are elicited. Each information structure is presented as a table with one row per signal ("clue"), which states the likelihood that the signal is realized and the likelihoods of a black and a white ball conditional on it (Figure 1(b) presents an example). The experiment has two parts. Part 1 consists of 15 guessing tasks, as in the main study, with the structures presented in the table format (Figure 35 in Appendix F). Part 2 consists of 80 choice tasks (Figure 36 in Appendix F shows the instructions). In each task, two information structures are displayed in random order, and the participant chooses which of the two they would rather receive information from before guessing the color of the randomly selected ball (Figure 37 in Appendix F). For each participant, the 80 pairs are drawn as follows: eight sets of seven structures are randomly selected without replacement from the 237 structures, subject to the constraint that each set includes structures spanning the full range of instrumental values, as in the ranking tasks of the main study, and ten of the 21 pairs of each set are randomly selected as choice tasks.

All participants received a base payment of \$6 and an extra \$0.5 if they answered all seven comprehension questions correctly on their first attempt.[58] A quarter of the participants were randomly selected to receive an additional bonus payment based on their decisions in one randomly selected part of the experiment. If Part 1 is selected, one of the 15 guessing tasks is randomly chosen and the participant receives \$10 if the guess in that task is correct. If Part 2 is selected, one of the 80 choice tasks is randomly chosen, and the participant receives information from the structure chosen in that task in a final guessing task and receives \$10 if the guess is correct. The median time participants spent on the experiment was about 23 minutes.

[58] The comprehension questions are essentially the same as those in the main study, with minor modifications to fit the tabular representation and the binary-choice design. Under the alternative design, 86 percent of participants answered all of them correctly on their first attempt.

E.2 Results

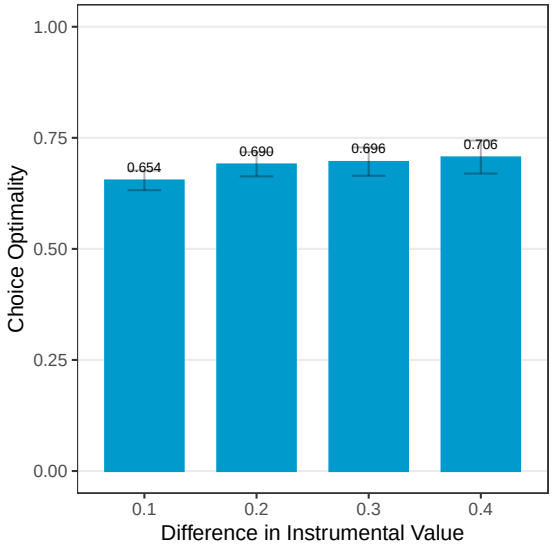


Figure 29: Choice Optimality by Value Difference under the Alternative Design *Notes: The figure plots the likelihood of the higher-value structure being chosen, by the difference in instrumental value between the two structures, in binary choices. Gray lines denote 95 percent confidence intervals by bootstrapping.*

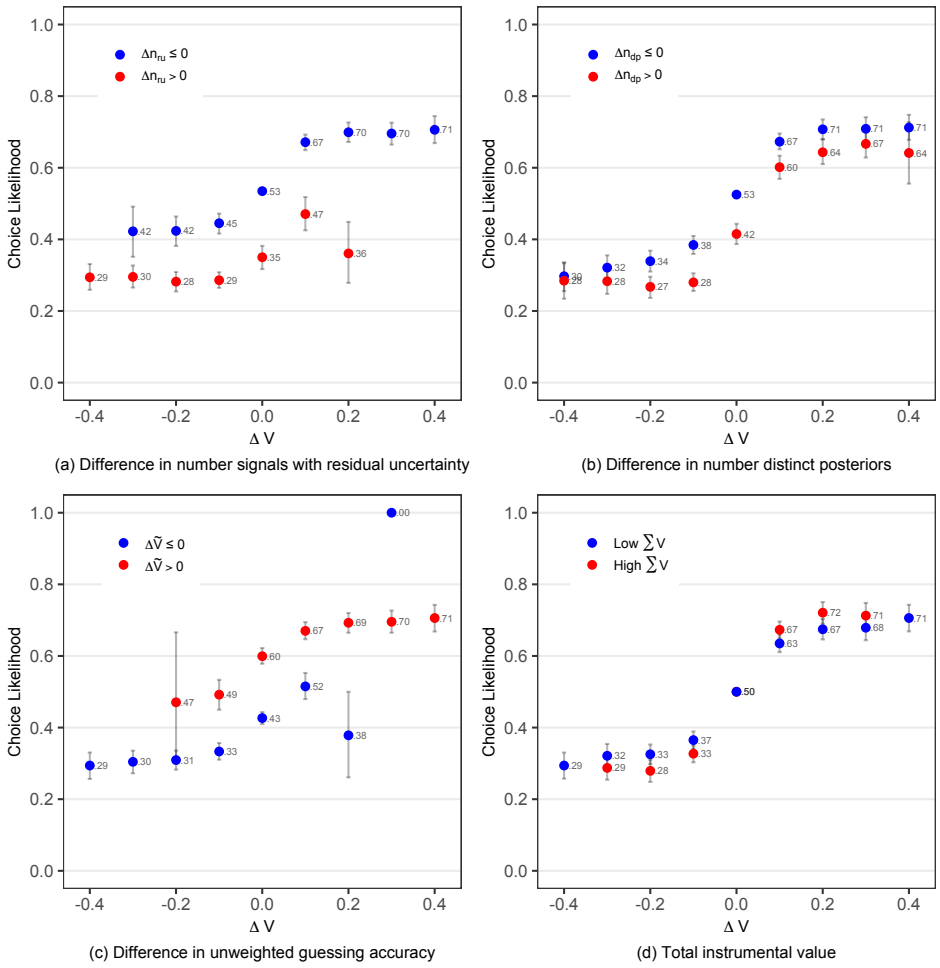


Figure 30: Choice of Information Structures by Value Difference under the Alternative Design: The Four Representative Characteristics of the Main Study *Notes: This replicates Figure 7 using data from the alternative design. Each figure plots the rate at which one structure was chosen over the other by value difference between the two structures and contrasts cases where values associated with a characteristic—difference in the number of signals with residual uncertainty, number of distinct posteriors, unweighted guessing accuracy, or total instrumental value—are either strictly above the median or weakly below. Due to symmetry, the median value is zero for the first three characteristics. Gray lines denote 95 percent confidence intervals by bootstrapping.*

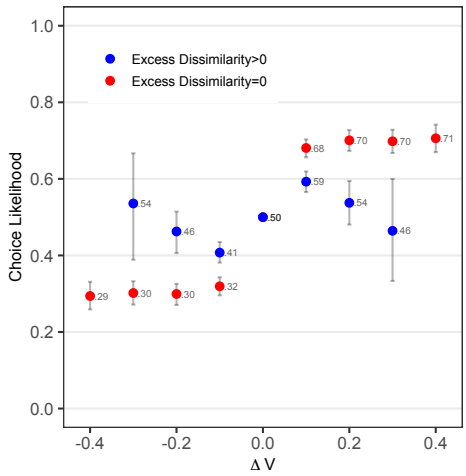
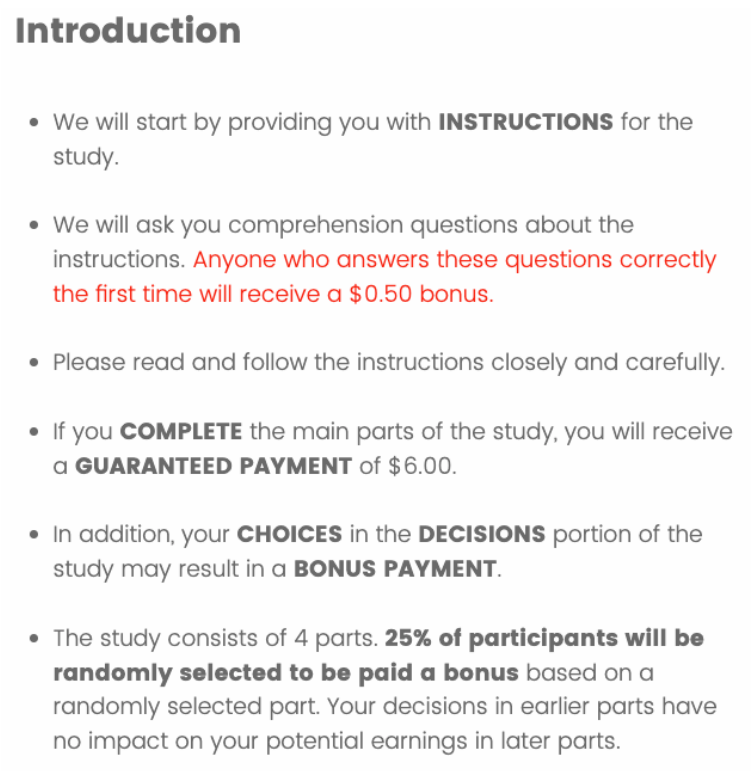


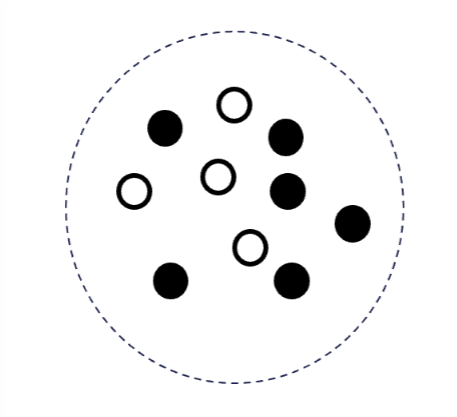
Figure 31: Choice of Information Structures by Value Difference under the Alternative Design: The Impact of Excess Dissimilarity *Notes: This replicates Figure 12 using data from the alternative design. The figure plots the rate at which one structure was chosen over the other by value difference between the two structures and contrasts cases where excess dissimilarity is strictly positive with cases where it is zero. Gray lines denote 95 percent confidence intervals by bootstrapping.*

F Instructions and Screenshots

The following are screenshots of the experiment software.



Part 1: Guessing task



- In each guessing task, there are 10 balls. 6 of them are black, and 4 of them are white.
- One of these balls has been randomly selected by the computer.
- Your task is to guess the color of the ball (black or white) the computer selected.
- You earn a $10 **BONUS** if your guess is correct and $0 if your guess is incorrect.

Each of the following comprehension questions was shown on a new page. Subjects had to answer each question correctly to proceed.

Comprehension question 1: Which statement is INCORRECT?

- ○ 6 of the 10 balls are black.
- ○ 4 of the 10 balls are white.
- ○ The chance that the randomly selected ball is black is 60 percent.
- ○ The chance that the randomly selected ball is white is 40 percent.
- ○ The chance that the randomly selected ball is white is 60 percent.

Comprehension question 2: Which statement is CORRECT about when you win a bonus of $10?

- ○ If your guess about the color of the randomly selected ball is correct.
- ○ If your guess about the color of the randomly selected ball is incorrect.
- ○ If you guess white.
- ○ If you guess black.

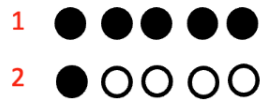
Information

Overview:

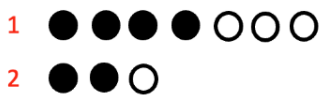
- Before you guess the color of the selected ball, you will receive some **INFORMATION** from an **INFORMATION SOURCE** about which ball the computer selected.
- Each information source divides the ten balls into different groups and informs you about which group the selected ball is in before you have to make a guess.

Description of an information source:

- An information source will divide the 10 balls into different groups like this, where **each row** is a different group and the group number is shown in red. Here is an example:



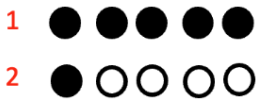
- and here is another:



(these are only examples -- we will show you many different information sources, with many different groupings across questions)

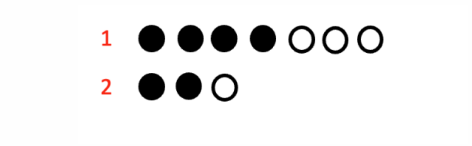
- We will tell you which Group number (shown as red numbers) the randomly selected ball is in **BEFORE** you guess the color of the ball.

For instance, in the following example



- The 10 balls are divided into two groups (two rows, labeled 1 and 2).
- There are 5 balls (all black) in group 1 and 5 balls (1 black and 4 white) in group 2.
- Since the groups are of equal size, the randomly selected ball is equally likely to be in group 1 or 2.
- If the information source reveals that the randomly selected ball is in group 1, then you will know for sure that the ball is black.
- If the information source reveals that the randomly selected ball is in group 2, then you will know that there is a 1/5 chance that it is black and 4/5 chance that it is white.

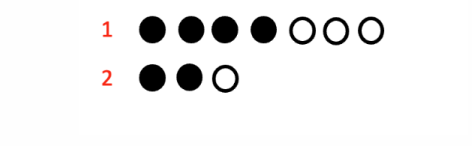
Comprehension question 1: Consider the following information source:



Which statement is CORRECT?

- ○ There is a 70 percent chance that the randomly selected ball is in group 1, and 30 percent chance that it is in group 2.
- ○ There is a 30 percent chance that the randomly selected ball is in group 1, and 70 percent chance that it is in group 2.
- ○ There is a 60 percent chance that the randomly selected ball is in group 1, and 40 percent chance that it is in group 2.
- ○ There is a 40 percent chance that the randomly selected ball is in group 1, and 60 percent chance that it is in group 2.
- ○ There is a 50 percent chance that the randomly selected ball is in group 1, and 50 percent chance that it is in group 2.

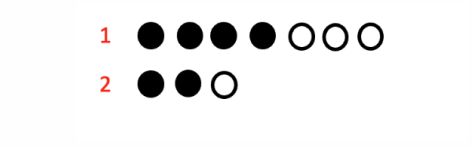
Comprehension question 2: Consider again the following information source:



If the information source reveals that the randomly selected ball is in group 1, what is the chance that is it black?

- ○ 4/7
- ○ 3/7
- ○ 1/3
- ○ 2/3
- ○ 1/2

Comprehension question 3: Consider again the following information source:

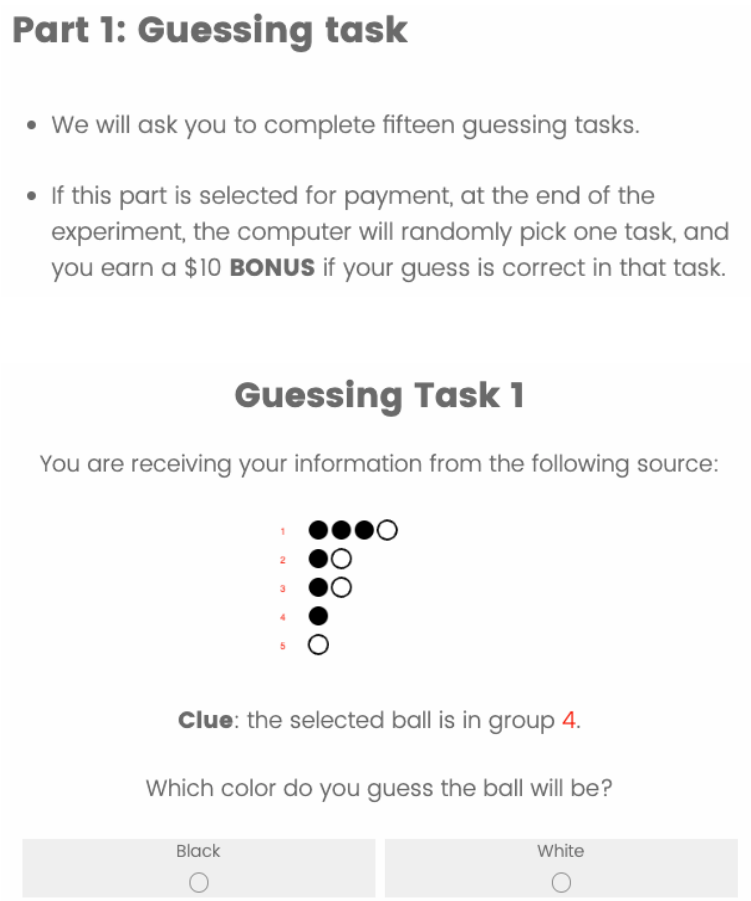


If the information source reveals that the randomly selected ball is in group 2, what is the chance that is it black?

- ○ 4/7
- ○ 3/7
- ○ 1/3
- ○ 2/3
- ○ 1/2

Comprehension question 4: In general, how do you maximize your chances of winning the $10 BONUS after receiving information from an information source?

- ○ Guess the color of the ball to be white.
- ○ Guess the color of the ball to be black.
- ○ Flip a coin to determine whether your guess is white or black.
- ○ Make your best guess about the color of the ball given the information provided to you on which group the selected ball belongs to by the information source.



This is an example of the guessing task. Subjects completed fifteen guessing tasks and then proceeded to part 2.

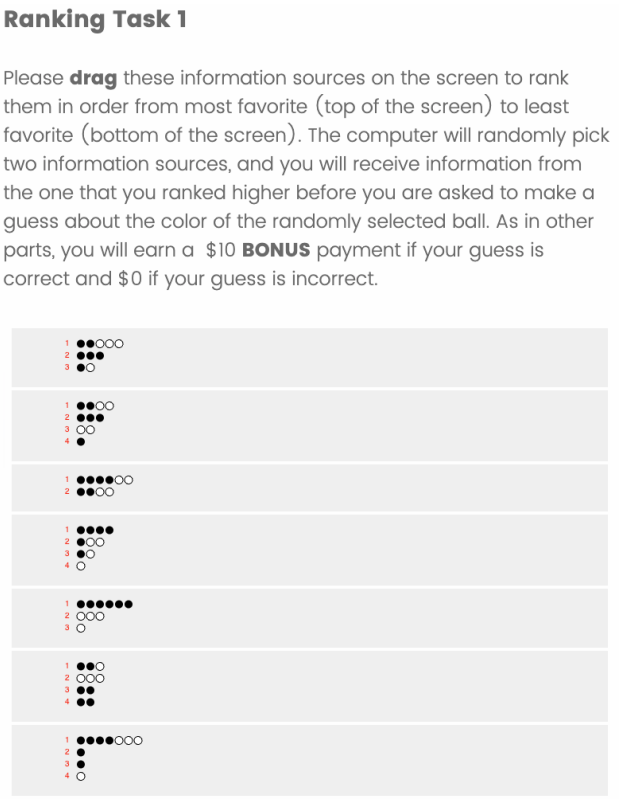
Part 2: Ranking information sources

Overview:

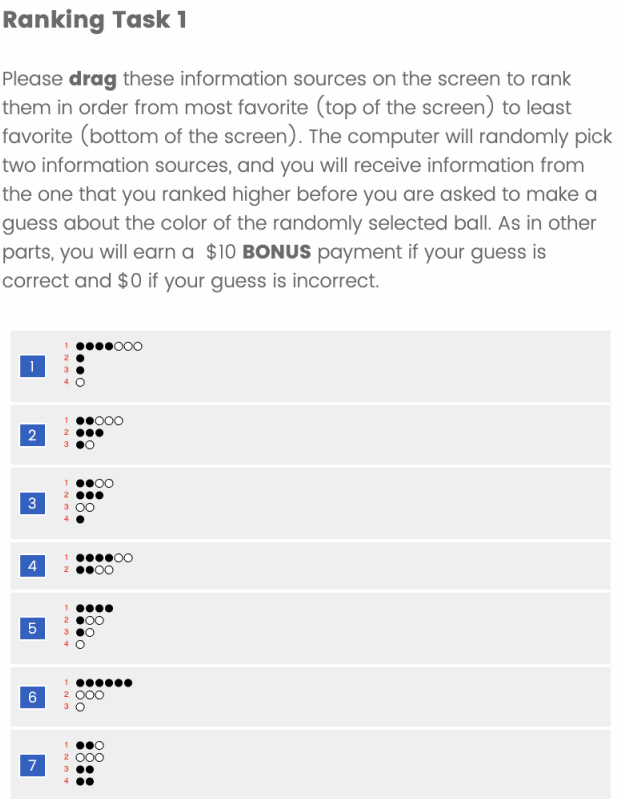
- You will tell us which information sources you would rather receive information from for a future guessing task by **RANKING** different information sources shown to you on your screen.
- Each information source divides the 10 balls (6 black and 4 white) into different groups as shown to you earlier.
- Your ranking (from highest to lowest) reveals which information sources you would rather receive information from before making a guess about the color of the selected ball.
- On each screen we will show you 7 different information sources in random order.
- You will use your mouse to click and drag these sources from most preferred (at the top) to least preferred (at the bottom).
- This part contains eight ranking tasks. If this part is selected for payment, at the end of the experiment, the computer will randomly select two of the information sources from one of the screens, and you will receive information from the one that you ranked higher before you are asked to guess the color of the randomly selected ball.
- Recall that you will earn a $10 **BONUS** payment if your guess is correct and $0 if your guess is incorrect, so your ranking choices could have a big impact on your earnings!

Comprehension question 1: Which statement is CORRECT?

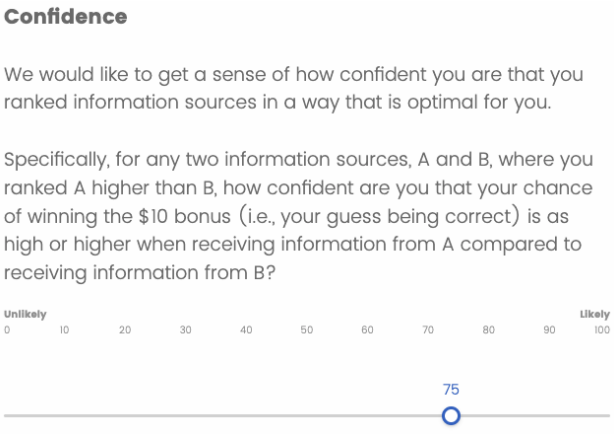
- ○ I am more likely to receive information from the information source I rank higher (above) than the one I rank lower (below).
- ○ I am more likely to receive information from the information source I rank lower (below) than the one I rank higher (above).
- ○ I am equally likely to receive information from any of the information sources.

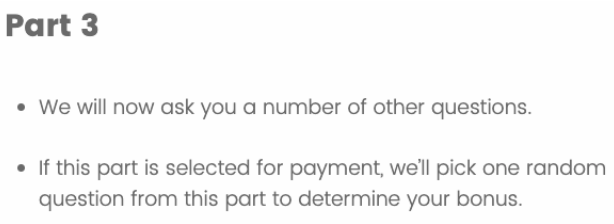


As subjects drag and reorder structures, an automatically updating order number appears to the left of each structure.

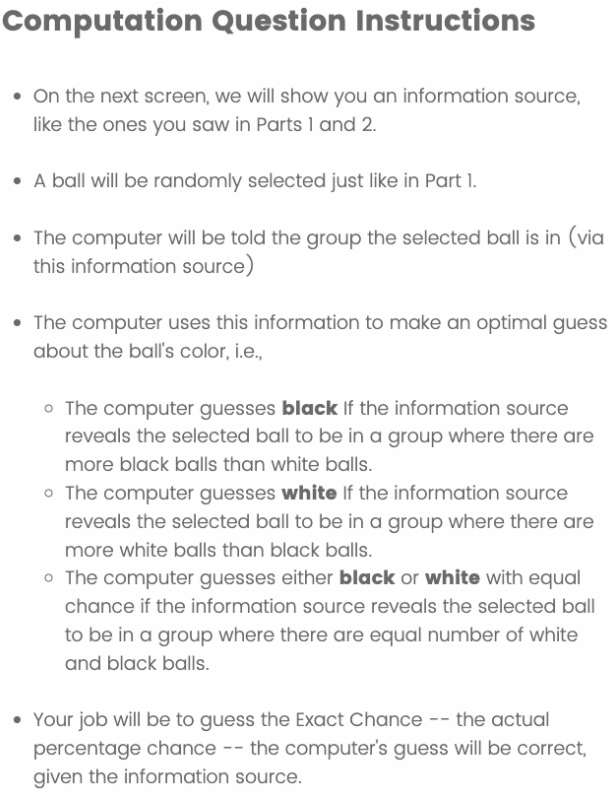


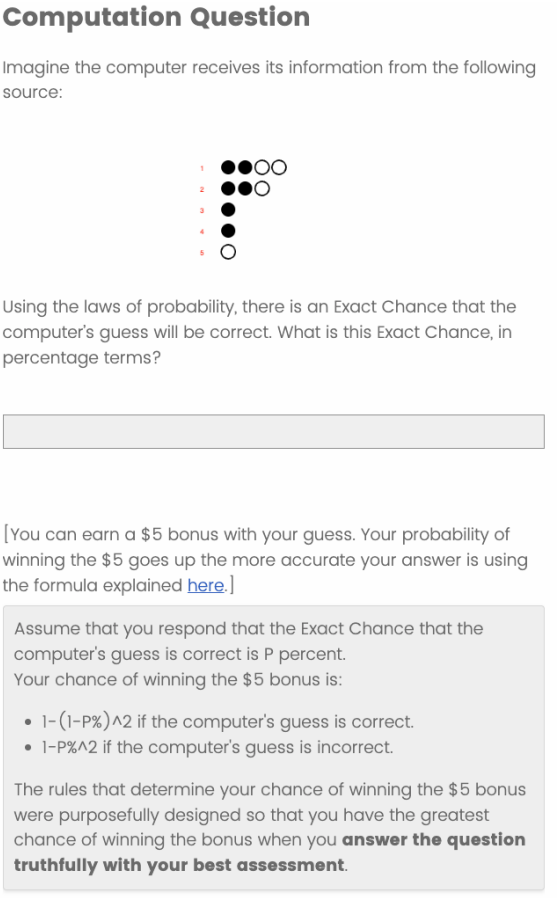
Subjects completed eight ranking tasks and then answered the following confidence question.





Part 3 contains several tasks. Subjects first completed the information structure computation task described below.

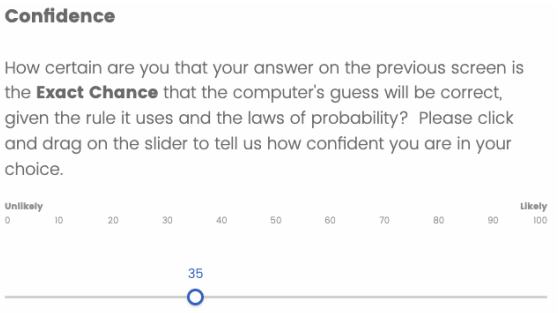




The detailed payoff calculation rule was initially hidden but would show up after subjects clicked the hypertext "here".

This shows one of the two pre-selected information structures used in this question. Another information structure is a partition having two cells (groups): group 1 contains six black and three white balls, and group 2 contains only a white ball.

After answering the above computation question, subjects reported their confidence on the next page.



Subjects then completed three tasks that elicited their certainty equivalents for three lotteries with equal expected value – a simple 50/50 lottery, a compound lottery, and an Ellsberg lottery. Below are the instructions.

Choice Task Instructions

On the next pages, you will be asked to choose between a **fixed amount of money received for sure** or **winning $30 if some event happens**.

For example, suppose there are a number of cards and each card is either Purple or Green. The computer shuffles the cards and then randomly draws one of them. You will be asked to make a choice between two options like below:

$30 if drawn card is Purple	○ ○	**$15** for sure

- If you select the **left option**, you win $30 if the randomly drawn card is Purple and $0 otherwise.
- If you select the **right option**, you win $15 for sure, regardless of the color of the randomly drawn card.

In fact, you will answer a list of questions like this. For example:

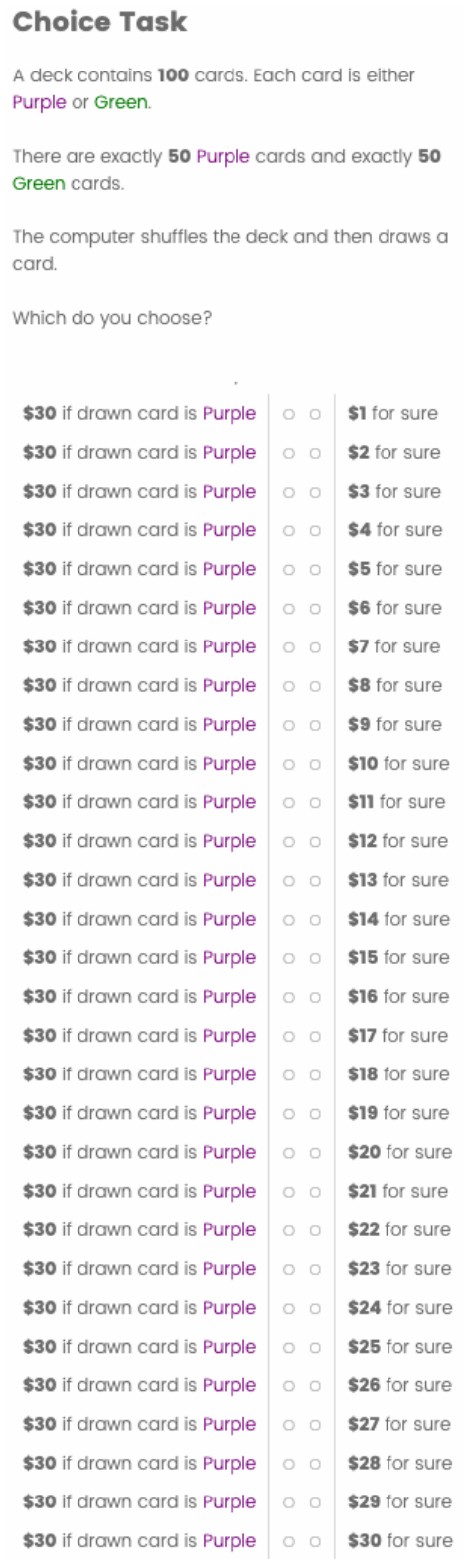
$30 if drawn card is Purple	○ ○	**$10** for sure
$30 if drawn card is Purple	○ ○	**$11** for sure
$30 if drawn card is Purple	○ ○	**$12** for sure
$30 if drawn card is Purple	○ ○	**$13** for sure
$30 if drawn card is Purple	○ ○	**$14** for sure
$30 if drawn card is Purple	○ ○	**$15** for sure
$30 if drawn card is Purple	○ ○	**$16** for sure
$30 if drawn card is Purple	○ ○	**$17** for sure
$30 if drawn card is Purple	○ ○	**$18** for sure
$30 if drawn card is Purple	○ ○	**$19** for sure
$30 if drawn card is Purple	○ ○	**$20** for sure

The option on **the left does not change**, while the option on **the right gets better as you go down the list**.

You must make a choice in all rows, but for simplicity, you only have to **click once on the row where you want to switch from left to right**.

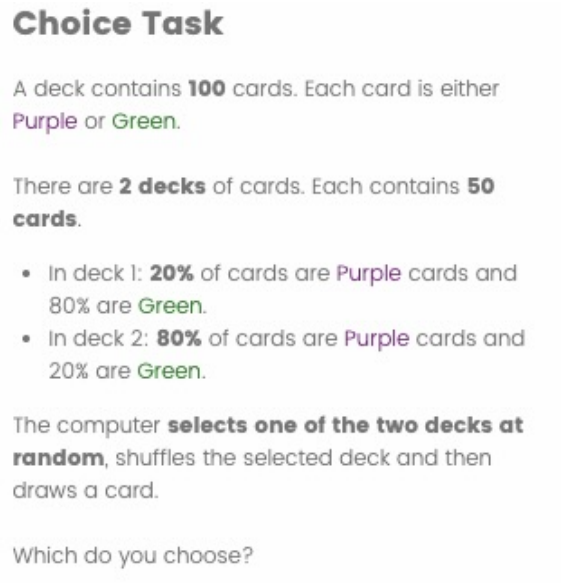
Intuitively, you can think about how much you would pay for a bet that would win you $30 if the randomly drawn card is Purple. Then, you switch to the sure amount of money as soon as it is above the amount you would pay for the bet.

At the end of the experiment, if this question is selected to determine your bonus payment, the computer will **randomly pick one row from the list** and pay you based on your choice in that row. This means you should **carefully consider your choice in each row** as any row could determine your bonus payment.

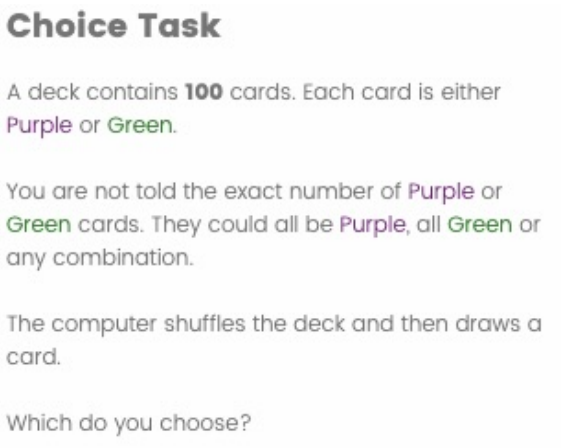


This is the elicitation of certainty equivalent for the simple 50/50 lottery. The other two elicitation tasks used the same multiple price lists. The order of the three elicitation tasks was randomized at the subject level.

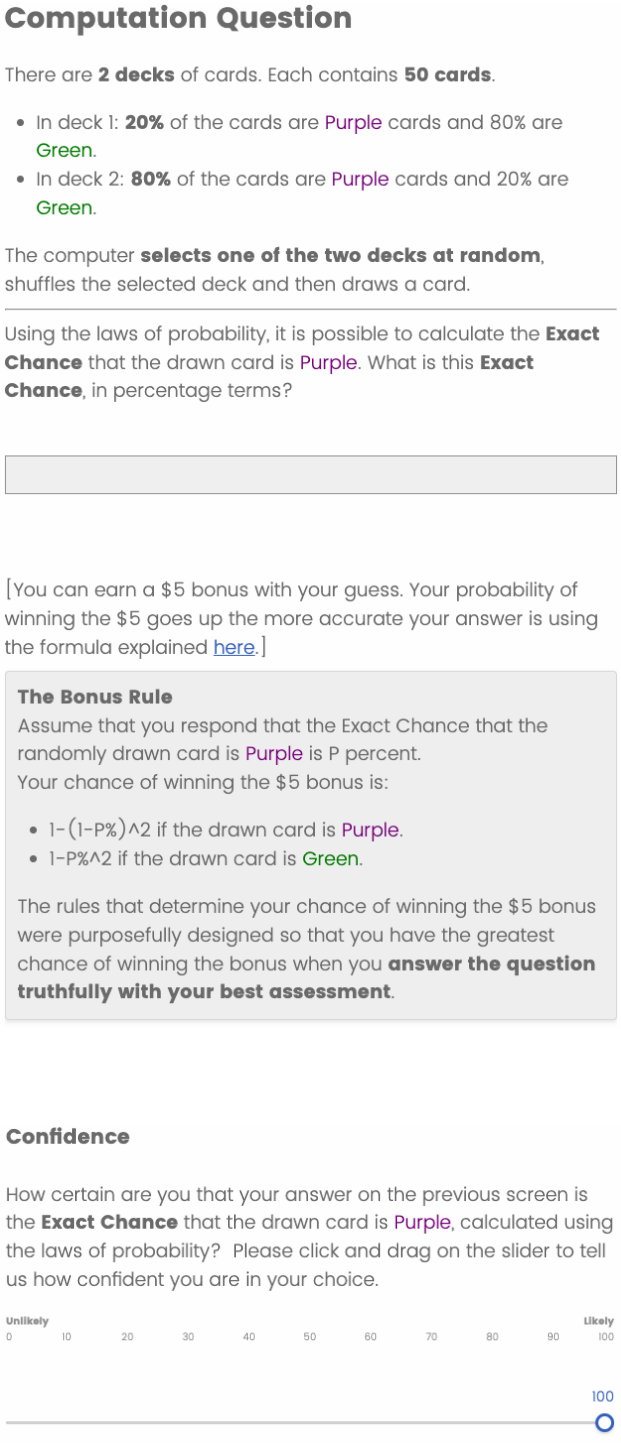
Compound lottery



Ellsberg lottery



Subjects then completed the following compound lottery computation question and reported their confidence on the next page.



Computation Question

There are **2 decks** of cards. Each contains **50 cards**.

- In deck 1: **20%** of the cards are Purple cards and 80% are Green.
- In deck 2: **80%** of the cards are Purple cards and 20% are Green.

The computer **selects one of the two decks at random**, shuffles the selected deck and then draws a card.

Using the laws of probability, it is possible to calculate the **Exact Chance** that the drawn card is Purple. What is this **Exact Chance**, in percentage terms?

[You can earn a $5 bonus with your guess. Your probability of winning the $5 goes up the more accurate your answer is using the formula explained here.]

The Bonus Rule
Assume that you respond that the Exact Chance that the randomly drawn card is Purple is P percent.
Your chance of winning the $5 bonus is:

- 1-(1-P%)^2 if the drawn card is Purple.
- 1-P%^2 if the drawn card is Green.

The rules that determine your chance of winning the $5 bonus were purposefully designed so that you have the greatest chance of winning the bonus when you **answer the question truthfully with your best assessment**.

Confidence

How certain are you that your answer on the previous screen is the **Exact Chance** that the drawn card is Purple, calculated using the laws of probability? Please click and drag on the slider to tell us how confident you are in your choice.

Unlikely 0 10 20 30 40 50 60 70 80 90 Likely 100

100

Part 4 contains ten cognitive questions, each displayed on a different page in the experiment.

Part 4

In this Part we are going to ask you a set of 10 questions. If this part is selected for payment, you will earn 50 cents for each correct answer.

Question 1

Below are four cards. Each has a symbol on one side and a number on the other side. You are given the following statement which may or may not be true of all four cards: "If a card has a star on one side, then it has a number larger than 10 on the other side."



Your task is to name those cards, and only those cards that need to be turned over in order to determine whether the statement is true or false.

- ☐ Star card
- ☐ Sun card
- ☐ 7 card
- ☐ 13 card

Question 2

A farmer had 15 sheep and all but 8 died. How many are left?

- ○ none
- ○ 7
- ○ 8
- ○ 10

Question 3

If you're running a race and you pass the person in second place, what place are you in?

- ○ first
- ○ second
- ○ third
- ○ fourth

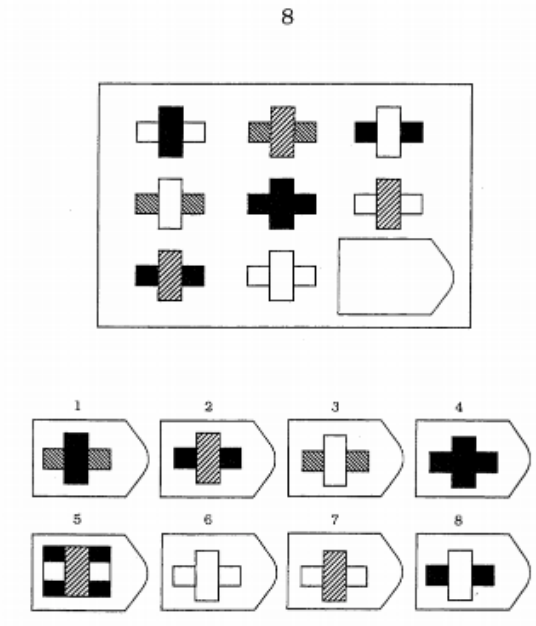
Question 4

Emily's father has three daughters. The first two are named April and May. What is the third daughter's name?

- ○ Can't know from information given
- ○ June
- ○ May
- ○ Emily

In each of the next three questions (5-7), there is a pattern with a piece missing and a number of pieces below the pattern. You have to choose which of the pieces below is the right one to complete the pattern.You will see 8 pieces that might complete the pattern. In every case, one and only one of these pieces is the right one to complete the pattern.

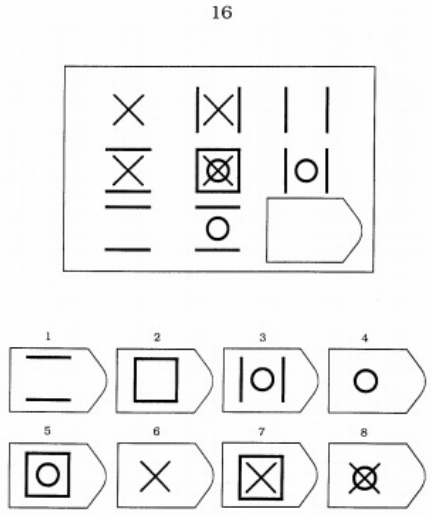
Question 5



Which piece completes the pattern?

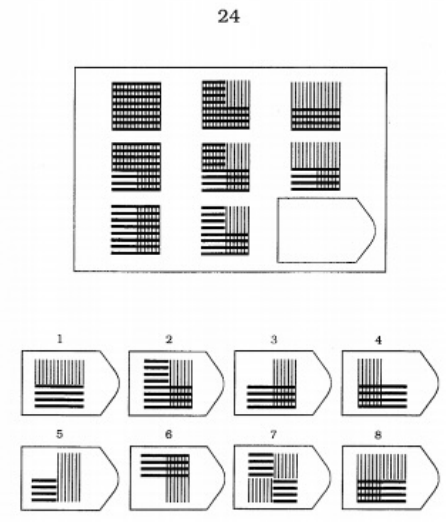
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7
- ○ 8

Question 6



Which piece completes the pattern?

Question 7



Which piece completes the pattern?

For each of the following three problems (8-10), decide if the given conclusion follows logically from the premises. Select YES if, and only if, you judge that the conclusion can be derived unequivocally from the given premises. Otherwise select NO.

Question 8

Premise 1: All flowers have petals.
Premise 2: Roses have petals.
Conclusion: Roses are flowers.

Does the conclusion follow from the premises?

- ○ YES
- ○ NO

Question 9

Premise 1: All mammals walk.
Premise 2: Whales are mammals.
Conclusion: Whales walk.

Does the conclusion follow from the premises?

- ○ YES
- ○ NO

Question 10

Premise 1: All animals like water.
Premise 2: Cats do not like water.
Conclusion: Cats are not animals.

Does the conclusion follow from the premises?

- ○ YES
- ○ NO

Guessing Question 1

There are 6 blue balls, and 4 red balls. One of these balls will be randomly selected and you earn $10 if you correctly guess the color (blue or red):



You will learn which of the following Groups the ball is in before you guess:



Please tell us what color you will guess if you learn the ball is in each of these possible Groups:

If I learn the ball is in **Group 1**, I will guess the color to be:	○ **Blue** ○ **Red**
If I learn the ball is in **Group 2**, I will guess the color to be:	○ **Blue** ○ **Red**
If I learn the ball is in **Group 3**, I will guess the color to be:	○ **Blue** ○ **Red**

(Remember,these choices determine your actual guess and therefore your payment!)

Figure 32: Guessing Task in Laboratory Study



(a) Ranking

(b) WTP

Figure 33: Ranking and WTP Elicitation Tasks in Laboratory Study

Ranking Information Sources

Please drag these information sources on the screen to rank them in order from most favorite (top of the screen) to least favorite (bottom of the screen). The computer will randomly pick two information sources, and you will receive information from the one that you ranked higher before you are asked to make a guess about the color of the randomly selected ball. As in other parts, you will earn a $10 BONUS payment if your guess is correct and $0 if your guess is incorrect.

(To help with your decision, for each Information Source, we include below each Group the guess you made for that Group earlier in a corresponding Guessing Question.)

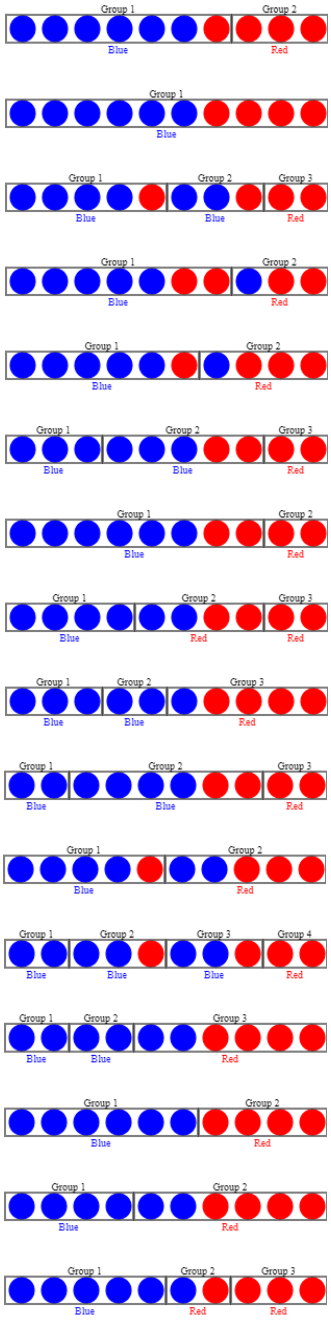


Figure 34: Ranking Task from Reverse Treatment (Guess First) in Laboratory Study

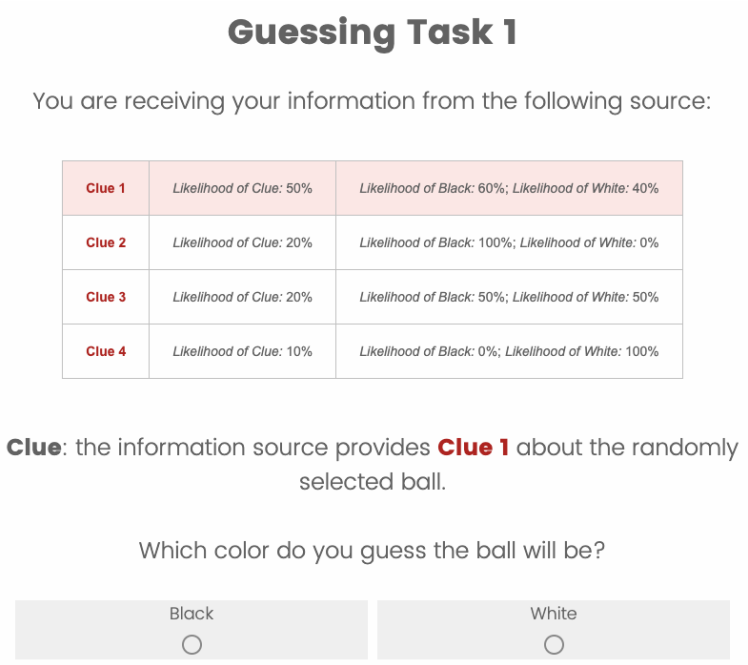


Figure 35: Guessing Task under the Alternative Design

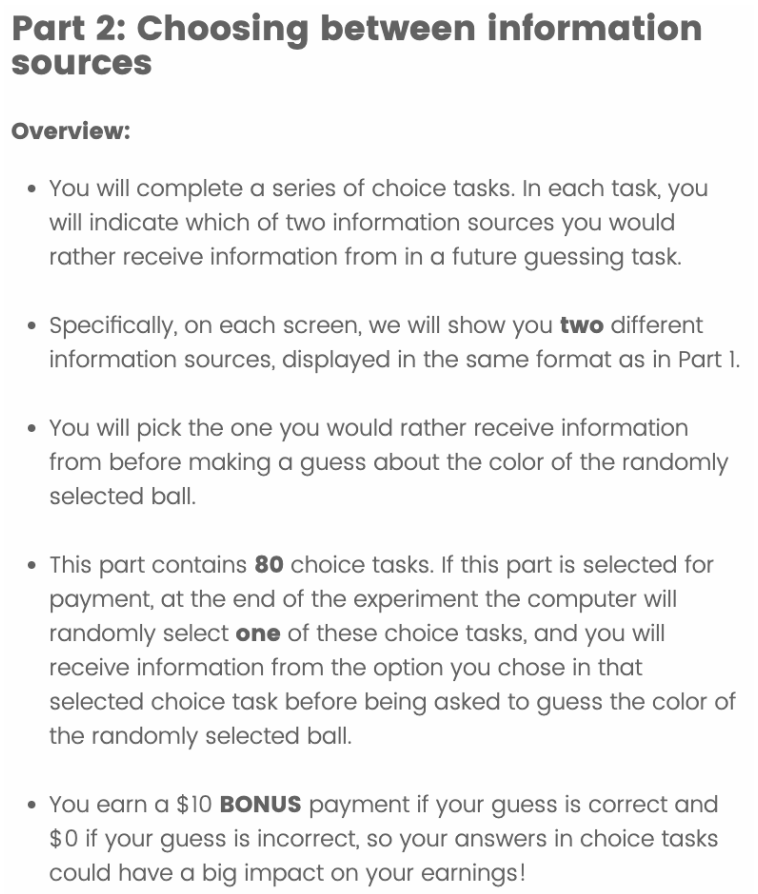


Figure 36: Instructions for the Choice Tasks under the Alternative Design

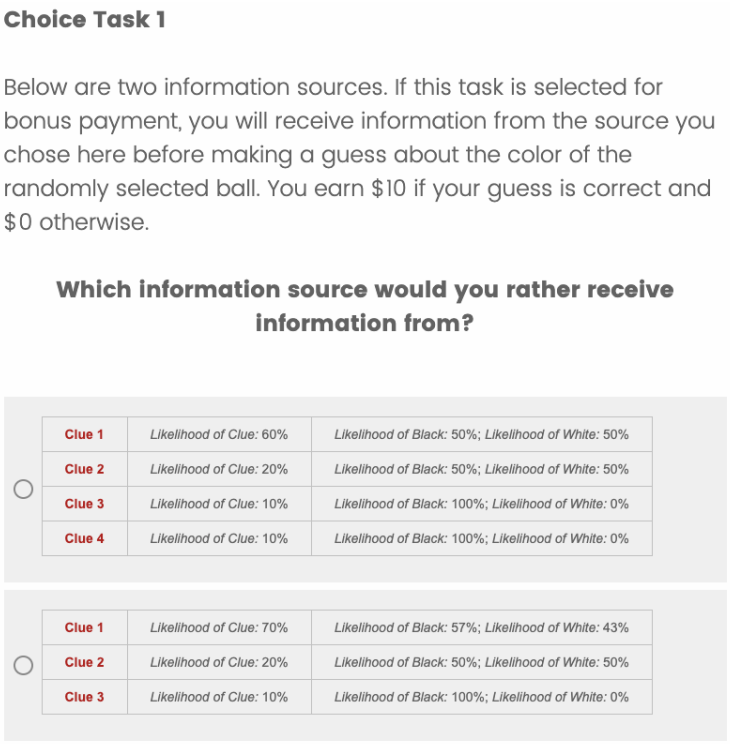


Figure 37: Choice Task under the Alternative Design